

1 CGGGCGCGGC GCGGCGGGCG GTGACAGCGG CGCCCGCGCC TCCCCGCGCG  
 51 TAGGTGTGCG GCGCGCTCCT GCGGAGGACG GAGCGAGCAG ATCTCGCGTG  
 101 CGCTCGCCGC CCGGCGCAGC CCAGCCCGGC CCCCCTGG CGCCGCGAGC  
 151 CGAGGTGTCT CCCGCGCCCG CGCCCGTGTG GCCGCGGTGC CCGCGAGCGG  
 201 GAGCCGGAGT CGCCGCCGCC CGAGCGCAGC CGAGCGCAGC CCGAGCCCGT  
 251 CCGCCGCCGC CATGGCCACC ACGGTGACCT GCACCCGCTT CACCGACGAG  
 301 TACCAGCTCT ACAGGATAT TGGCAAGGGG GCTTCTCTG TGGTCCGACG  
 351 CTGTGTCAAG CTCTGCACCG GCCATGAGTA TGCAGCCAAG ATCATCAACA  
 401 CCAAGAAGCT GTCAGCCAGA GATCACCAGA AGCTGGAGAG AGAGGCTCGG  
 451 ATCTGCCGCC TTCTGAAGCA TTCCAACATC GTGCGTCTCC ACGACAGCAT  
 501 CTCCGAGGAG GGCTTCCACT ACCTGGTCTT CGATCTGGTC ACTGGTGGGG  
 551 AGCTCTTTGA AGACATTGTG GCGAGAGAGT ACTACAGCGA GGCTGATGCC  
 601 AGTCACTGTA TCCAGCAGAT CCTGGAGGCC GTTCTCCATT GTCACCAAT  
 651 GGGGGTCGTC CACAGAGACC TCAAGCCGGA GAACCTGCTT CTGGCCAGCA  
 701 AGTGCAAAGG GGCTGCAGTG AAGCTGGCAG ACTTCGGCCT AGCTATCGAG  
 751 CTCAGGGGG ACCAGCAGGC ATGGTTTGGT TTCGCTGGCA CACCAGGCTA  
 801 CCTGTCCCT GAGGTCCTTC GCAAAGAGGC GTATGGCAAG CCTGTGGACA  
 851 TCTGGGCATG TGGGGTGATC CTGTACATCC TGCTCGTGGG CTACCCACCC  
 901 TTCTGGGACG AGGACCAGCA CAAGCTGTAC CAGCAGATCA AGGCTGGTGC  
 951 CTATGACTTC CCGTCCCTTG AGTGGGACAC CGTCACTCCT GAAGCCAAAA  
 1001 ACCTCATCAA CCAGATGCTG ACCATCAACC CTGCCAAGCG CATCACAGCC  
 1051 CATGAGGCC TGAAGCACC GTGGGTCTGC CAACGCTCCA CGGTAGCATC  
 1101 CATGATGCAC AGACAGGAGA CTGTGGAGTG TCTGAAAAAG TTCAATGCCA  
 1151 GGAGAAAGCT CAAGGGAGCC ATCCTCACCA CCATGCTGGC CACACGGAAT  
 1201 TTCTCAGTGG GCAGACAGAC CACCGCTCCG GCCACAATGT CCACCGCGGC  
 1251 CTCCGGCACC ACCATGGGGC TGGTGGAACA AGCCAAGAGT TTACTIONACA  
 1301 AGAAAGCAGA TGGAGTCAAG CCCAGACGA ATAGCACCAA AAACAGTGCA  
 1351 GCCGCCACCA GCCCAAAGG GACGCTTCTT CCTGCCGCC TGGAGCCTCA  
 1401 AACCACCGTC ATCCATAACC CAGTGGACGG GATTAAGGAG TCTTCTGACA  
 1451 GTGCCAATAC CACCATAGAG GATGAAGACG CTAAAGCCCG GAAGCAGGAG  
 1501 ATCATTAAGA CCACGGAGCA GCTCATCGAG GCCGTCAACA ACGGTGACTT  
 1551 TGAGGCCTAC GCATTCTACT TCGAGAACCT GCTGGCCAAG AACAGCAAGC  
 1601 CGATCCACAC GACCATCTTG AACCACACG TGCACGTCAT TGGAGAGGAT  
 1651 GCCGCTGCA TCGCTTACAT CCGGCTCACG CAGTACATTG ACGGGCAGGG  
 1701 CCGGCCCCG ACCAGCCAGT CTGAGGAGAC CCGCGTGTGG CACCGCCGCG  
 1751 ACGGCAAGTG GCAGAACGTG CACTTCCACT GCTCGGGCGC GCCTGTGGCC  
 1801 CCGCTGCAGT GAAGCCAAGG GAGGGGCACA GAATGGGGAA CAGGACACAG  
 1851 GATCCTAAAC TCCAAGGGGA CTGTCCACCG ATGAACACTC AGAGTGGACA  
 1901 CCATCTTCCG TCCACGCTGT GCCCAGGACA GCTGTCCCA TCCATGAACA  
 1951 CAGGGTAAAC ATCTGCCGGG CTCCGCACCA GTGGCTCCCT GGGCCATGGG  
 2001 ACAGCGGCAC GGCTCACCAC GGACAGCAGC TGGCCCAGCA GCCGGCCACC  
 2051 CTGGCGTCCT GGGGCTCCT CCCCTCCTCT CCCTCTCACC TTGTACCTC  
 2101 CACGGAGCTG CCTGTCTGGG ATAATTGGG GATTTTTTTT TCTGGGGGAT  
 2151 AATTCTTTTG CATGACCCCT AAAGAGCAAG CCACACCGGT CTGCTAGCTA  
 2201 GGTGTCCGCG GTGTGGTG (SEQ ID NO:1)

**FEATURES:**

5'UTR: 1-261  
 Start Codon: 262  
 Stop Codon: 1810  
 3'UTR: 1813

# Homologous proteins:

## Top 10 BLAST Hits

			Score	E
CRA	18000005245285	/altid=gi 5326757 /def=gb AAD42035.1 AF07880...	1047	0.0
CRA	18000005199792	/altid=gi 10835006 /def=ref NP_001211.1  cal...	1044	0.0
CRA	18000004938668	/altid=gi 6671660 /def=ref NP_031621.1  calc...	1039	0.0
CRA	18000004937301	/altid=gi 11120682 /def=ref NP_068507.1  Ca+...	1038	0.0
CRA	18000005245287	/altid=gi 5326762 /def=gb AAD42037.1 AF08192...	1001	0.0
CRA	18000005171302	/altid=gi 3668373 /def=gb AAC79460.1  (AF085...	999	0.0
CRA	1000737074531	/altid=gi 6688228 /def=emb CAB65122.1  (AJ252...	986	0.0
CRA	18000005245288	/altid=gi 5326764 /def=gb AAD42038.1 AF08341...	986	0.0
CRA	18000004964693	/altid=gi 466360 /def=gb AAA81938.1  (U06636...	982	0.0
CRA	18000005199791	/altid=gi 4139268 /def=gb AAD03743.1  (AF112...	982	0.0

## BLAST dbEST hits:

	Score	E
gi 12801212 /dataset=dbest /taxon=960...	1675	0.0
gi 12868201 /dataset=dbest /taxon=960...	1453	0.0
gi 2053138 /dataset=dbest /taxon=9606 ...	1247	0.0
gi 10213950 /dataset=dbest /taxon=96...	1243	0.0
gi 9324431 /dataset=dbest /taxon=960...	1233	0.0
gi 12921378 /dataset=dbest /taxon=960...	910	0.0

## EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|12801212 Fetal brain

gi|12868201 Fetal brain

gi|2053138 Testis

gi|10213950 Lung small cell carcinoma

gi|9324431 uterus endometrium adenocarcinoma cell libe

gi|12921378 Fetal brain

Tissue expression from PCR-based tissue screening panels:

hippocampus

1 MATTVTCTRF TDEYQLYEDI GKGAFSVVR CVKLCTGHEY AAKIINTKKL  
 51 SARDHQKLER EARICRLKH SNIVRLHDSI SEEGFHYLVF DLVTGGELFE  
 101 DIVAREYYSE ADASHCIQOI LEAVLHCHQM GVVHRDLKPE NLLLASKCKG  
 151 AAVKLADFGL AIEVQGDQQA WFGFAGTPGY LSPEVLRKEA YGKPVDIWAC  
 201 GVILYILLVG YPPFWDEDQH KLYQQIKAGA YDFPSPEWDT VTPEAKNLIN  
 251 QMLTINPAKR ITAHEALKHP WVCQRSTVAS MMHRQETVEC LKKFNARRKL  
 301 KGAILTTMLA TRNFSVGRQT TAPATMSTAA SGTTMGLVEQ AKSLLNKKAD  
 351 GVKPQTNSTK NSAAATSPKG TLPPAALEPQ TTVIHNPVDG IKESSDSANT  
 401 TIEDEDAKAR KQEIIKTTEQ LIEAVNNGDF EAYAFYFENL LAKNSKPIHT  
 451 TILNPHVHVI GEDAACIAYI RLTQYIDGQG RPRTSQSEET RVWHRRDGKW  
 501 QNVVHFHCSGA PVAPLQ (SEQ ID NO:2)

# **FEATURES:**

## **Functional domains and key regions:**

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
 N-glycosylation site

Number of matches: 3

1 313-316 NFSV  
 2 357-360 NSTK  
 3 399-402 NTTI

[2] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE  
 CAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 2

1 48-51 KKLS  
 2 259-262 KRIT

[3] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
 Protein kinase C phosphorylation site

Number of matches: 4

1 47-49 TTK  
 2 51-53 SAR  
 3 358-360 STK  
 4 367-369 SPK

[4] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
 Casein kinase II phosphorylation site

Number of matches: 9

1 36-39 TGHE  
 2 51-54 SARD  
 3 79-82 SISE  
 4 94-97 TGGE  
 5 109-112 SEAD  
 6 262-265 TAHE  
 7 400-403 TTIE  
 8 401-404 TIED  
 9 485-488 SQSE

[5] PDOC00007 PS00007 TYR\_PHOSPHO\_SITE  
 Tyrosine kinase phosphorylation site

9-17 RFTDEYQLY

[6] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 3

1 302-307 GAILTT  
2 332-337 GTTMGL  
3 390-395 GIKESS

[7] PDOC00100 PS00107 PROTEIN\_KINASE\_ATP  
Protein kinases ATP-binding region signature

20-43 IGKGAFSVVRRRCVKLCTGHEYAAK

[8] PDOC00100 PS00108 PROTEIN\_KINASE\_ST  
Serine/Threonine protein kinases active-site signature

132-144 VVHRDLKPENLLL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	195	215	1.665	Certain
2	319	339	1.301	Certain

# BLAST Alignment to Top Hit:

```
>CRA|18000005245285 /altid=gi|5326757 /def=gb|AAD42035.1|AF078803_1
      (AF078803) calcium/calmodulin-dependent protein kinase II
      beta subunit; CAM2 [Homo sapiens] /org=Homo sapiens
      /taxon=9606 /dataset=nraa /length=542
      Length = 542
```

```
Score = 1047 bits (2678), Expect = 0.0
Identities = 516/542 (95%), Positives = 516/542 (95%), Gaps = 26/542 (4%)
Frame = +1
```

```
Query: 1      MATTVTCTRFTDEYQLYEDIGKGAFSVVRRCVKLCTGHEYAAKIINTKKLSARDHQKLER 180
      MATTVTCTRFTDEYQLYEDIGKGAFSVVRRCVKLCTGHEYAAKIINTKKLSARDHQKLER
Sbjct: 1      MATTVTCTRFTDEYQLYEDIGKGAFSVVRRCVKLCTGHEYAAKIINTKKLSARDHQKLER 60

Query: 181    EARICRLLKHSNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQI 360
      EARICRLLKHSNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQI
Sbjct: 61    EARICRLLKHSNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQI 120

Query: 361    LEAVLHCHQMGGVVRDLKPENLLLASKCKGAAVKLADFLAIEVQGDQQAQWFGAGTPGY 540
      LEAVLHCHQMGGVVRDLKPENLLLASKCKGAAVKLADFLAIEVQGDQQAQWFGAGTPGY
Sbjct: 121    LEAVLHCHQMGGVVRDLKPENLLLASKCKGAAVKLADFLAIEVQGDQQAQWFGAGTPGY 180

Query: 541    LSPEVLRKEAYGKPVDIWACGVILYILLVGYPFFWDEDQHKLYQQIKAGAYDFPSPEWDT 720
      LSPEVLRKEAYGKPVDIWACGVILYILLVGYPFFWDEDQHKLYQQIKAGAYDFPSPEWDT
Sbjct: 181    LSPEVLRKEAYGKPVDIWACGVILYILLVGYPFFWDEDQHKLYQQIKAGAYDFPSPEWDT 240

Query: 721    VTPEAKNLIQNMLTINPAKRITAHEALKHPWVCQRSTVASMMHRQETVECLKKFNARRKL 900
      VTPEAKNLIQNMLTINPAKRITAHEALKHPWVCQRSTVASMMHRQETVECLKKFNARRKL
Sbjct: 241    VTPEAKNLIQNMLTINPAKRITAHEALKHPWVCQRSTVASMMHRQETVECLKKFNARRKL 300

Query: 901    KGAILTTMLATRNFSVGRQTTAPATMSTAASGTTMGLVEQAKSLNKKADGVKPQTNSTK 1080
      KGAILTTMLATRNFSVGRQTTAPATMSTAASGTTMGLVEQAKSLNKKADGVKPQTNSTK
Sbjct: 301    KGAILTTMLATRNFSVGRQTTAPATMSTAASGTTMGLVEQAKSLNKKADGVKPQTNSTK 360

Query: 1081   NSAAATSPKGTLPAALEPQTTVIHNPVDGIKESSDSANTTIEDEDAKARKQEIIKTTEQ 1260
      NSAAATSPKGTLPAALEPQTTVIHNPVDGIKESSDSANTTIEDEDAKARKQEIIKTTEQ
Sbjct: 361   NSAAATSPKGTLPAALEPQTTVIHNPVDGIKESSDSANTTIEDEDAKARKQEIIKTTEQ 420

Query: 1261   LIEAVNNGDFEAYA-----FYFENLLAKNSKPIHTTILN 1362
      LIEAVNNGDFEAYA                      FYFENLLAKNSKPIHTTILN
Sbjct: 421   LIEAVNNGDFEAYAKICDPLTSFEPEALGNLVEGMDFHRFYFENLLAKNSKPIHTTILN 480

Query: 1363   PHVHVIGEDAACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRDGKWQNVHFCSGAPVAP 1542
      PHVHVIGEDAACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRDGKWQNVHFCSGAPVAP
Sbjct: 481   PHVHVIGEDAACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRDGKWQNVHFCSGAPVAP 540

Query: 1543   LQ 1542
      LQ
Sbjct: 541   LQ 542 (SEQ ID NO:4)
```

# **Hmmer search results (Pfam):**

Model	Description	Score	E-value	N
PF00069	Eukaryotic protein kinase domain	306.2	3.9e-88	1
CE00022	CE00022 MAGUK_subfamily_d	293.8	1.3e-86	1
CE00359	E00359 bone_morphogenetic_protein_receptor	15.0	0.0015	1
CE00031	CE00031 VEGFR	0.9	2.1	1
CE00287	CE00287 PTK_Eph_orphan_receptor	-65.4	0.00046	1
CE00292	CE00292 PTK_membrane_span	-77.0	0.00018	1
CE00291	CE00291 PTK_fgf_receptor	-93.1	0.0021	1
CE00286	E00286 PTK_EGF_receptor	-132.2	0.0059	1
CE00290	CE00290 PTK_Trk_family	-161.3	0.00033	1
CE00016	CE00016 GSK_glycogen_synthase_kinase	-196.7	9.2e-06	1

## **Parsed for domains:**

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
CE00359	1/1	132	186 ..	272	327 ..	15.0	0.0015
CE00031	1/1	133	205 ..	1068	1139 ..	0.9	2.1
CE00286	1/1	14	252 ..	1	263 []	-132.2	0.0059
CE00290	1/1	15	253 ..	1	282 []	-161.3	0.00033
CE00291	1/1	14	267 ..	1	285 []	-93.1	0.0021
CE00292	1/1	14	267 ..	1	288 []	-77.0	0.00018
CE00287	1/1	14	270 ..	1	260 []	-65.4	0.00046
PF00069	1/1	14	272 ..	1	278 []	306.2	3.9e-88
CE00022	1/1	10	305 ..	13	316 ..	293.8	1.3e-86
CE00016	1/1	1	343 [.	1	433 []	-196.7	9.2e-06

```

1 GAGCTGCTGT GTCTCTGTCC CCAGGGGCAG AGGGGCTGTG GGGTTGCAGG
51 CTCAGCGTCT GGGACTCTGG GGTGAAGGCT CAGCCATGCC CTGCAGACAC
101 CATGGGGCAG GGCTCAGACC TGTGCACCTG TCTCTTGCAA ACCACTGTTT
151 TCTCTGTTTT GTAACCCCCC ACCCAACCCC ACATAACACC TCTGGGTTTA
201 AACAACATGC ACCCTTGTGC CGGTACCTC CCTGCAGCCG GAGAACCTGC
251 TTCTGGCCAG CAAGTGCAAA GGGGCTGCAG TGAAGCTGGC AGACTTCGGC
301 CTAGCTATCG AGGTGCAGGG GGACCAGCAG GCATGGTTTG GTGAGTGCCA
351 GGGGCAGGGT GTGTTGGCTG GCAGTTGGCA GGGCAGGAGG TGATGCTGAC
401 AGCCCCTTGT GGCCTCTTCC CCTCTCTCTA GGTTTCGCTG GCACACCAGG
451 CTACCTGTCC CCTGAGGTCC TTCGCAAAGA GGCCTATGGC AAGCCTGTGG
501 ACATCTGGGC ATGTGGTGAG GCCTGGCCTG AGTTGGTGCG GGGCAGGGCC
551 TCGGGTGTTC CAGGACTTCC CACCTACATC CTGGAGTGTG CAGTGGCCAG
601 CACGTCTTGC TCTCATCTGG GTTTATCTGT GTCAGACCTG CCCTTGAGCT
651 GCCCTGGCAG GGGTCTGCCC ACACAGCCAA GAGCCCCCTT TCCACCAGA
701 TTAGAATTGC TCACATGAAC CTGGCGCACC CCAGTGCTCG CCTGCGCTCA
751 GCAGAGGTCT GGTCCAGAAG TGTGGTGGGT GGATGGGAGT GGAGAAGAGA
801 GGTCAGGGGC TGTGGGGCCA TGGGCAGGGC CACCTCCTTG GGTAGGGGTC
851 TCCTCCCACA GAGGTGGGGA GCAGCAGAGG GGCTTGACAT CACCCTCATC
901 CCTGTGATAG TGTGGGTGTG GGGCAGAGGT CAGGGGGCCG GCTGTGCCCT
951 TCTACCCAG TGTCTGCTGC ACAGGTGGGG GCAAAGGAAT GCTGAGGACC
1001 CCAATGCCCT CCCAGGGCCA CAGGAGCTAG GCAGTGAGGG TGCAGGGCAT
1051 GGGCTTCATG GACGGTGGCA CCCTGCAAGT GGCTGCGGTG CTCACAGGCC
1101 CCATCCGCAG GGGTGATCCT GTACATCCTG CTCGTGGGCT ACCCACCCTT
1151 CTGGGACGAG GACCAGCACA AGCTGTACCA GCAGATCAAG GCTGGTGCCT
1201 ATGACGTGAG TGCACCAGCC CCTCTCTGAT GAGCTCCCTT CCTCCAGGTG
1251 TGGCCGGGTG AGGGCAGCGT GGGGAAGAGG TAGGAGTGGG GTGAAGCCAC
1301 CTGTGGCCAG GTCCTGGGTC CTGCTCTCCC AGATTCTGGG CTGGAGATGA
1351 AGCCCCTTGG AGAATTCTTG CCCCTGCCTG AGAGGGAGCT TCAGGCCCGG
1401 CCGGGGCGCT GTTTCCTTCT GCAGTTCCCG TCCCCTGAGT GGGACACCGT
1451 CACTCCTGAA GCCAAAAACC TCATCAACCA GATGCTGACC ATCAACCCTG
1501 CCAAGCGCAT CACAGCCCAT GAGGCCCTGA AGCACCCTGT GGTCTGCGTG
1551 AGTCGCCCTT GGTGCCCATG GTGGGGAGGG GGCTCCTGGT GGAGATGGCC
1601 TCAGACCACT CCCCTGGCAA GGACCCCAAG AGGGTCTGTG TCCTGACATC
1651 CAAGAGCTCC CTGGGTGCC CTGGGTGCTC CTTGTGGCCT CTGGCTTGGG
1701 ACATACCAGC ACGTTTGTGA GGCTTGGGCG TTGGAAGGCA TTAGAGGGTA
1751 GAGGTGATCC CTTCCTCCCA ACTGCAGTCC TGTCTGTGAG GGGCAGAGTG
1801 GACGAGGCAA GGGAGAGACG AGTCTTGAAG TCCAGGCGG GTGGGGACAG
1851 ACAACCCTTG CCGCAATGGT GGCCGGTGGC TCTTGGCAAG TGGGGACCCC
1901 AGGTGCCAC AAGCCTTGCC ACCCTGGCCT CTCCCCTGTG CCTCGGGCTC
1951 GGCTGCCATA TGACCACCCA TTTCCCCACA GCAACGCTCC ACGGTAGCAT
2001 CCATGATGCA CAGCAGGAG ACTGTGGAGT GTCTGAAAAA GTTCAATGCC
2051 AGGAGAAAGC TCAAGGTGAG GCCCTGGCCC CTAGTCCAG GCACGGCCAT
2101 GCTTCTCTGT GTCCCTCTGG GCTGGAGCAG GGGGGCCTTG GGGGTCTGG
2151 GCAGACCTAG GGGTTACTGC TGCCCCAAG ACTGACTGTT AGCAAGTCCC
2201 AGACTGGATG CATCAGGTGA ACTCAGGCCA GCTTGGGAAT GAGTCCAGAG
2251 GGGCCCTGGG CCAGGTGTGG CTCCTCCTAG TTGTCTGTGC CACCTCCTAG
2301 CAGCCCTTGG AGGAGCTGTC CTGAAGCGCT CGCTGTGGGC TCCTACCCG
2351 GGCTCTGCAG GCAGCACTCA CCTCTGGCA GTCACACTGT TTAGTACAAG
2401 CAAGTCCGAA GCTTCCGGCT CAGACAGGTT TGGTAAGGAG AGCAGAGCCA
2451 CACACACTGG TCTTGGGTGG GCTGGGGGAG TTCTGGGAGG GAGGTGGGTC
2501 CCAGTAGGGT ATCCAACCTG CCTGCTTTGG TCAGGGCTGG CTCCGGTGAC
2551 CGCACACTGG CAGTCCCTCT ACTTGTGGGT TCCGGGATGG GGACTTGTG
2601 CCTGACTGCC CTCTGCTGGT CTCTGAGCAG TTCTCCCCGG AAGCCCCAGG
2651 ACTGTTGCCG TGTCTGAGCC TGTGAGGAAA AGAAGGGGCT GTCAGGGAGC
2701 TGGACCCAG AGGAGCTGCC GTGGTGACCA GCTGTTCTGG TGACCCCTGA
2751 GGCTTGAGGG GTCTTGAAGC AGCTAGAAGC TGTAAGTTGGT CAACAGGTTT
2801 AGGCCAGGG TGTGTGTAGT TCTGGAAATA GGTGATCTGT CTCAGTGCAG
2851 CTGCTGGCTT CCTGGAGCTC TTGCCTCTCT GGAAGGCTGA GGTCATGTCA
2901 GCCTCATGAC AATGAGGCTG AGCATCTGGG CAGGAGGACA GGGGTCTTAT
2951 CCTGGCCAGA AGCCAGCAGG GAACACTGAT GGGATAGCCC CGGTTTTATC
3001 TGTGTCTCTC CCCAGGGAGC CATCCTCACC ACCATGCTGG CCACACGGAA
3051 TTTCTCAGGT GAGCCTTTCT TCTCCAGGGA GACAGGCGCT GCCCCCTCCC
3101 TGCTGGCCCA CGCAGGAGAG CGCCTCCTTC CTCACCAGCC TCTCCACTCC

```

FIGURE 3, page 1 of 16

3151 TCCTCTGCGG CAGGCCTGCC CTCGGCGTCT GCCCTCAGCT CTGAGACCCA  
3201 CTGCCCACCT GGCCCCGCTG GGCTCCCACC TTGGGTGATA CCACAGGGTC  
3251 CAGCCCCCGG AGGCCATCAC CTTCTGTGTG GGTCTGTGTC CCTCCACCCC  
3301 CTGAACACGA GCGTCTGTGC TGCCCCACTG GGGCTCACAG CATCGTGTGT  
3351 GTCTGTCCAG GCGTTTGTGC GGCATCTATG TGGCCTCCTT GTCATTTTGA  
3401 GTGCTCTGAA CATTGTGTTT TGTGCGGGAG GTGGGCAGAA GGGATGCGGG  
3451 GTGATGCGGG AGGCTCGGGG GCCTCCTTCC AAGTTCTGGA TGAGCTGCAG  
3501 CCTCCTGTCC CGGCTGCTCA GGGTGGGTGG TTGGGAAGCA AGTTCTCTTG  
3551 GCAGGGGGGT GGGTCTGTG ATAGACCCCT GAGGCCCAGG GCGCTGGCAG  
3601 ACCCATCGGG GCATGATGTT AGCCCCGGAG TGGAGCCGGC AGCCAGGTC  
3651 TGGACAAGCT GTACCTGTGG CTTCTCCGTC GTCCGACACT CCGTGTGCGA  
3701 GCGTCTGTGA TCCGTCTCTC TCGTTGTCCG TTTGCATCTG GTGCCCCCA  
3751 CCCGCCATCC TGTTACTTTT GCTGTGATGC TGTAATGCCG GGAACGCGTG  
3801 CACACGGTCA CACCAACACT AATAGGACTG TCCTGTCTGC TGTGTGCTCA  
3851 CCACACCCTT TGGGCATGAG AAGCCCCAC TGGGGTTTTT TAAGGAGAAA  
3901 GGAGGCAAAT GCTTTTCCGT GTCAATCAGT CCAATCTTGT TTCACTCTC  
3951 TTGAGCAAAG GATTCTGGAA CCATCTGTCA CCTAAACTTT AACTCTAATC  
4001 TTCTTCTGCT TCCTTTGTCT CTTTCTTCC CTTACCTCGC CCACCCCTCG  
4051 TCTGTGTCCG CCCACCCCTC CCTTCCCTC GTCTCTAACC CGGTGCTAAC  
4101 AGTGGGACAG CAGACCACCG CTCCGGCCAC AATGTCCACC GCGCCTCCG  
4151 GCACCACCAT GGGGCTGGTG GAACAAGGTA GATGTGTCTC GACCAGCGTC  
4201 CCGCCCGCTC CCGCCCGTCC CTCCTGCCAG CATGCAGCCC CCTGTGTCAC  
4251 GCAGCGCTG GCGCGGCTCC AGAGCCGCC CAGAGCGCGC CAGGCCCCCG  
4301 GGAGCCCCTG CTCCCGTGTG GTCACATCCC AGCAGAGCCC ACCACAAGGG  
4351 CAGGGAGGCA GCCCCAAGG CTCCTCGCCT GTAAGAGGAG GGGCTGGGCT  
4401 AGGTGGCCCC TGGGCTACAC CAAGCCCTT TGGTCTTGGC CCCCAGGTC  
4451 TGGGGTCCG GAGACCCCA TTAAGAATGG CCTGGGCCCC ACAGGGAGCC  
4501 ACTGGGCTG CTGCTGGGGG GTCTGAATCC TGAAAGGAGA GCCTTGAGGA  
4551 GCAGAGCCAG AGAGGCAGAG GCCCTTGGGG CAGACACACA CCCTGCCCTC  
4601 CTGGGGCCGC ATGGAGACGG TGGTCTGTGC TGCTGAGTCC TACACATGCA  
4651 TGTCTGCCCT GAGCATCCCC CCAGGACAAG CCGCTCTGGA GTGGGTGAGG  
4701 GTTTTATGCA CCCTGAGGAG ACTTTCAAGG CTTCTCTTG GGTGTTTCT  
4751 GCAAAGTCCT CCTCCCTGG CCTCAAACCC TGTGAGGGAA AAGGCCGCA  
4801 CTGGCCACCT GCTCCTCTGG GCTGTGCGG GCCAGAGCCC AGAGGCCCAA  
4851 GTTGGCTTCT GCCCACTGC TGGCTTGTGA CCATGGGCAG ACCCATGAG  
4901 GGCTAGGCGA CCCCAGACC TCCTTGAGC TCCAGCCTGA GCTGAAGGCT  
4951 GGTGAGAGCT TAGGGCAGGC CAAGCTGACA ACGCCTGGCC ACAGAACACA  
5001 GAGGGCTACA GGGGTGACCC CAGATCCTCC CTGGGCTGAG CTGCTGAGTT  
5051 CCCTGTGCGT GCCTCCAACG TGGGCTGGGG ACCCGGCAGA GGTTCCAGGG  
5101 TGCTGGAGAC TGCCTTCCCC AGGCCTCCTC ATGACCCACA GGGTGAGCAG  
5151 CCTGGCCTTC CCAGCCAGAG AACCTCCTT CTGGGGAGGC CCAGGGCGTC  
5201 CTCGGGGAGG GCAGTCTATT CTCCTCCCAT GAGCCAGTG GACGTGTCTA  
5251 GCAGGCAGCA CCCCAGGAGA GCCCTCCAC GTCTTCTCCA TTTGACAGGC  
5301 CTTTCCAGAG CGCAGGCGGG AGGGGGCTGT GATTAGAAAA GAGTGAGGCT  
5351 AGTGGCTTCT GGGGAGGCAC TGCTGCCAG GGGACAGTGC TGAGAGACAG  
5401 CTGCCTCTAC GCTGCCCTGT GCCCGGGCT CCCGCTGCAA TGCCCGCCTG  
5451 TCTGCAAGTG AACGTGGGGC GACGGTGCAT GAGGCCCTGC ATGTGTGGCT  
5501 CCACCCTGGG CGCCGAGAGC AGCTCTGTCC TGGAGGGTGG TCAGTGCATG  
5551 TGGACAGAGC CCAGCATGGC TGTCCTGGGT GACCAGCTAA GGGGACAAGG  
5601 CAGAGGCAGG GCTGAGAGGA CCACCCATCC TGCTAGGTCA GCCCAGCTCA  
5651 GCCATATCAC ACGGCAGTGA GCATGGAGCT CAGTTCTCTG CCAATGGCAG  
5701 CTGAGTCTAG TACCATCCAG TCAGAGTCTG GTACCAGCCC ATGTGGCATA  
5751 GCCCCCTCGG CCCGACAGAG GACCCCGTCT GTCGAGTGTG CTTAGTTTG  
5801 GCCTCTGTGG TCTCTCTGC ATTGATCAGG TGTAAGGGCA TAGGAGACCC  
5851 AGTGTCCGGC CAGCTGCAGG GTGGCAGCAG TTGCCCCGGC CTGGAGACCC  
5901 GGGAATGGGC AGTGCCTTCC CAGGATGGAG GGCAGAGGGT CTCTCCTTGT  
5951 CCCACAGAGG CCTGCAGAAC CCCCACCCA GGTGTCTGAG ATGCCTGTGA  
6001 CTGCTCCGCC TACCCTGGGC TCCTGCGGCA CCTAACGCAT GCTTTGAACT  
6051 TGAGACACAG AAAGGAAGTT CCCGTGCCCT TGAATGCTAG TGTAGATGGG  
6101 CATCGACAGG ACTCTGGCCA CGGTGAATCT GGAGTTAGTC CCAGGCAGAG  
6151 ATGTGAAATG AGCAGCCCCC CAAAAAATGG TTGGCCGGGA GCCATGCACT  
6201 CAGGAGGGCC GGGCCCATGC ACCCCACACT GCGCCCAAGG CGTGCACAAG  
6251 CGATTGTTTT AAAAGCGGGT TCACAAGGAA GGATGTTTGG GAACTGACTG

FIGURE 3, page 2 of 16



6301	AGACAACAGG	GACGTCTGCT	GCAGGGCTTC	CCAGAGCTCT	GATGGCAGCG
6351	TCGGCCTGAG	TCCTTCGAGG	AGGGCTGGTT	TGTACGTGGC	ATTGCTGCC
6401	CACCTGGACTG	TGAACCTCTG	TCTTTTATT	TCCCACTGCT	GCTGTGGTAC
6451	ATCTCCAGTA	GCATAGTTTG	GAAATGCAGG	TTTTGATAGA	CTCAAGGATC
6501	TAAATAGAAC	CCTCTTAGTA	CCAAGGACTG	TCCGGGGTCT	CTGCCAGCCC
6551	CGCCGATGGG	CCTAACTGTG	GTGCCTCCTT	TCCTGTGAGA	ATCTTCTGAG
6601	GACATGCCCG	GGGAAAGAGC	TCAGTTCTGC	TGCTGCCTAG	GGTGCCATGC
6651	TGGCCCCGGT	TCCAATGCAG	AGCCTAGCTG	GAAGTACCGC	TGGGTGGCG
6701	GAGGCTACGT	GCCTGACTGT	CCCCTCGGGG	GTGGGGTGGA	ACTAGCCTTC
6751	TGAAACCGCC	TGCTTCAGTT	GGCCACAGCT	TTTTGAAATG	TGTGTTTCTG
6801	GAAGGGACTG	GGTCCCTTCC	TTGCCTGTTC	AGCTCCCCAC	GACAAATGTC
6851	CTCAAGGCGA	GGCTGGATGC	TTCTTCCTC	AGGCTCCTAG	GAGGAGCCCG
6901	TCCCCCAGCT	GTGTCGGGCA	GCTGGTCACC	AGCAAGGACA	GGATCCCTCA
6951	GCTGCAGCCT	CAGGCTGGCT	GGCACTGGGC	GGGTGTTTCT	GGGATGAGTT
7001	GTGTGTACTG	GAGTAGGGAG	GGGAGCTGAG	AGGGTGGGAT	GCACAGACAG
7051	GAGAGGGGAC	TGTGGGGGTC	CTGGAACCTT	GAGTTCCAAG	TCTTCAGGAC
7101	TCTCCCTCCA	TAGCAAGTTA	CAGGGAAGCA	GATTTGAGCC	ACAGGGAAGC
7151	AGATTTGAGC	TGCAGCGAGG	GGGAGGGTTT	TCAGTCTGTG	CTATAGGGAA
7201	GTGGGCAGTC	GGCATTTCCTG	GTCCCTGGGAA	CTCACTGGGC	AGGGCTGCCT
7251	TGGGACATCA	GGGAGGTGCG	GCTGTGCTCA	GCTTCACCAG	GAGGGGCCCTT
7301	AGGCCTGGGG	ACGGAGAGTG	ATGCCTGAGG	CCCCTCTACT	TCTCCATGGA
7351	TCCTGGGAGG	GACTCCTGGG	CTGGATACAA	AATTGTTGAG	AGTTAAGAGA
7401	TCTGTGAGGA	AGGGGAGGCT	GGGAATAGAA	AGTGTGTGCC	CACTGCACAT
7451	GGGGTCCGCA	GGGCCACGTG	CAGCCACTGC	GCAGGCACAA	CCCCAGTCCC
7501	CACAGAGCCC	AGGAGGGGCC	AGAGCCATGG	AGGAGGCAGC	ACTGGGCATT
7551	TGGACAGGGA	GGGGGTGGTC	AGCAGGCAGC	AGGCCCAGGC	CTGTCTATGC
7601	CCTGCGGGGT	GCAGCCTCCT	GATCTCCACG	GCAACCTGGA	GCACCCAGCG
7651	TCAGAACCAC	CGGGAGGGCT	TATGGAACAG	ATGTCCAGCC	CTGCAGAAGT
7701	TCTGGCTCAG	GAGGGCGGGG	TGGGCCTGGG	AATTTGCATT	TCTGACTGTA
7751	CAGGGCGATT	CTGCTGCTGC	TGCTGCTGCT	GGGGTTGGGG	GAGGATCCCA
7801	TTTGAGAAGC	GCTGCAGTCC	TAGGTTGAAA	CGTGCCTGTC	TGTCCCCACC
7851	CAGGCCTGCA	TGGGCAGCAC	GGGATCCCCA	GGCAGGAGGA	CCCAATTTCA
7901	TGGCCTGGCC	AGCCAGGGTC	CTGGAGCCAG	GCGGTGGGGG	AGGGATGGGG
7951	GATTGCTGTG	CCACCTTCCT	TCCCGGCTTG	GCCCGGGGCG	AAGCATCCTC
8001	ACACTTCCCA	TGTCGTATC	CCCTTGGCTC	CAGCCTGGCT	GCCTCTCTAA
8051	CCCTGCTGTA	CCGGCTGGCC	GCATGGCCCT	GGCTCTTTT	GGTGAGCGTG
8101	GTCCAGGACT	GGTGACCTGT	GAGTCTGGG	CCCGCAGTCT	TGCGCCCCTG
8151	CCCGAACCAG	CACAAATCTT	GTTTTCTCTC	TCTCTCTTCC	TTCTCTACTC
8201	CCTCCCTTTC	TCACCTTTTC	TTTTCTGTAA	GGTAAGCTGA	CTTCTCTTTT
8251	TGGTTTTTTA	TTTATTTTAA	TTTTTTAGTT	CTGTAATTAA	AATCCTAACA
8301	GCCATTGGAG	GTGTGGGCAC	CGGGGGCTGG	GGCCAGGCCC	CTCTGACCTC
8351	TGAGGGGGAA	TGCTGGGTGA	GGCAGGGGCC	CCGCTGCTGG	GACCAAGTAT
8401	CCTCAGGGGC	TTGTGGGCAG	AAAGGCCTGT	GCTGGCCCCA	GTCAGTGCAC
8451	AGAAGCGGCC	CCAAGGCCAG	GGCTGCTGGG	CAGCTCGGAA	TGAGGGCGAG
8501	CAGGGCTGCC	CTTGGTGCCT	GAGCCAAGGA	GCCAATGGGA	CAGACCTCTG
8551	AGCCTGGGTG	CCAAGTATGA	GGTCTGAGAC	AGGGTGAGCG	CCTGGGCTGG
8601	GACAAGGCCC	TCTGAGTGGG	CGGCCAGCTG	CAGCCCACCC	ACCCCTACCC
8651	CAGGAAGGCA	GGGCCCGGGA	GGGCATGACC	TCTGGGGTGC	TGGCTCAGCT
8701	GCCCCACCC	CAACCTGACA	CCGCTAGTCC	TGAGTTCCCA	TCAGGGAGGA
8751	AGCAGCATCC	TGCCTTCCTC	TAGGAAGAGC	TTGCATGTGG	CCCAGAAGCC
8801	AAGGGGGCTC	CCCAGCACCC	ACGGGCATCT	CTGGGTCTGG	TCAGAGGAGA
8851	AATCTGGATG	CTTGCAAGGAG	CCCCAGGGTC	ATGGAGGAGG	CTGGAGACAG
8901	GGCTGTCTTG	GGGTGATGGG	ATGGCCCCCC	CACCTGCTCA	GAGCCAGCCT
8951	GGGTGCTGGA	ACCACACTTG	CCTCAGGACC	CTGGGCTTGC	TCCTGGGGAA
9001	AGAGTGGGGT	CAGGCAAAGG	GGTGGGGTTG	CGCTGCAGCG	AGACCCAGGC
9051	CCATCACTCA	CCATACCTTC	TTCTTCCCCA	TGCAGCAGCC	AAGAGTTTAC
9101	TCAACAAGAA	AGCAGATGGA	GTCAAGGTGA	GGCTCCAGCC	GGGCCCTGTG
9151	GTGCCGGGGA	GCCAGAGGCC	TGCAGCTTCA	CCCCACGCC	CTGGGGCTCC
9201	TGCTCTGGAG	TCCCCCTCCC	CCCATGCCCT	GAGAGACACG	GGACAGGGAA
9251	TGGCGAGTGA	GGGGCTTCTC	CCACCTAAGA	GTTCCTCTTC	CCTCTCTCCA
9301	CAGCCCCAGA	CGAATAGCAC	CAAAAACAGT	GCAGCCGCCA	CCAGCCCCAA
9351	AGGGACGCTT	CCTCTGCGG	CCCTGGTACT	GAGCTCCTCA	AATTCTGCCT
9401	CTCAGCCCC	CCTACGCCCC	TGGCTGTGTG	ATTGCCGCTG	GTCAGAGGGG

FIGURE 3, page 3 of 16

9451	GCCGGGTGAA	GGTGGGGTCT	GGCCCCGCCT	GGCTGTCTG	ACAGCACTCG
9501	CATGGCCCC	GCCCCTCATC	CCTCACCGGT	GGTGAAGTGG	AGAGAAGAGG
9551	CCACTGTTGT	GGGGGGCTCC	AATTCAGACA	GGTTTAGGAC	TGCTCTGGGG
9601	AGCCCCCTGGC	TGAGACCCAC	AGATGTTGGG	GTGCAGGGGA	GAGGCCACGC
9651	CTCCCACCCA	TGTTGACTTG	TGGATGTCTC	TCCAGGAGTG	TTCAGGAAGT
9701	CAGTGAGGCA	GAAATACCC	TCTCCCCACC	AGGACCCAC	CCTCAGCTCC
9751	TCCACCATCC	TCAACAGGCC	GACCCACAGA	CCACTCCGAA	GGTCTGGCTT
9801	GGTGGGGCTG	GGCCAGGATC	TGCAGGGGGA	ACAGCCCATA	GTGGCACATT
9851	CCACGGCCCA	TGGGGAGACG	GGGCCACGGT	GGTGCAGTAG	AGAGGTGTCT
9901	AAGCCAGTGG	CAGCCAAGGG	GAGGGCTTGC	CGTCACCTCT	GTGTTCCCTC
9951	AGTGCTGTCT	TGTGGCTGCC	TGAGAGGCAG	GGCTTAGGGG	CTCCCTGCCG
10001	GGGAGGGGAG	GGGTCCCCAC	CATGCTCCGC	TCCAACCTGC	CCCCCTAGTG
10051	CCCTTGTCCC	TGGGGGCTCC	TACAGTGTAA	CCCATATAGCA	GTACTCCCAA
10101	GGATGTAAAG	TTGTGGCTGG	TGGGTGCCGG	CCTTCCTGCT	GGGGCGCTGT
10151	GCTGTGTCCC	CTCAGCTGTC	CTAAGAGCTT	TGGGGCTTGC	TGGCCCCGTAG
10201	GTCCCCATAT	TTGCTGGAAG	CAGGCTTGGT	GTCCCCTGAG	AACCCCAGGC
10251	CAGGCTTCGG	GAGCCAGCCC	CAGACCGCCC	ACGGGAATAC	TGGGTTTGCC
10301	AAATGGCCAC	CTTGAGACCC	AGGAGAGGAG	AGCGGTCTCT	GGAGGGCCGA
10351	CTGTGCTCAGA	GCAGCCAGGC	CGTGGCTGGA	GGGTGBCCTG	GTGCAGCCTA
10401	CCTTAGGGCCT	CTCAGTGGCC	AGGGCAGCCC	ACGTGCCAGC	CTCACAGCCA
10451	GCCCCATCTC	GGACCCTGTC	CATCCCATGT	GCCACCGCCA	CCCCCATGAC
10501	ATCTTCAAAC	CTGTGCCCCC	CACCACGCTG	GGGCACAGGT	TCAGGCAGTA
10551	AAGGGTAGGG	AGAACCCCTC	AAGACCGAGC	CTGGCTTCTC	TGGCTCCCCAC
10601	ACACATTGTG	CAGCTTTGTCG	GGGCCCCACA	CGGTCCATCT	CCCACCCTGG
10651	ACACGAGCAC	CTCCGCCAGC	CTGGACAGAG	CTCCTGTCCA	TTCCATCCCT
10701	CCCGGCTGAC	CCAGGCTCCT	CCCCCAGCTG	CTCCACGCCG	CCTCCATCCC
10751	TGTCCCCCAC	TCTGCTCTGC	ACTTCTTTCT	CGCAGGCTCT	GGCCACCCAC
10801	ACCTCCTCTG	TCTCCCTGTT	CCCCTCCTGG	TGGTCTCCGC	TTCTCCTCT
10851	TCTCACTTTC	CCTCTCTTTC	CTTCCTCTGT	GTCTTCCTTC	TTCTGTAGGA
10901	GCCTCAAACC	ACCGTCATCC	ATAACCCAGT	GGACGGGATT	AAGGTACTGC
10951	CCAGCTTTCC	TCCTCCCGTT	TTCCCCAGGC	AGGAGGCTCC	AGGCCAGGAG
11001	AGAGGTCTGG	GCAGCATTTT	GTGCCAGAGT	GGAGGGCAGA	TGTCCTATGG
11051	CCCTGGCCGC	CCCTCCCCGC	AGTACGGTAG	GGCCCCAGTC	CGTCTTCGTG
11101	GGCAACAACA	GGACAGACTG	GCTCAGGCCC	CAGGCGCGCC	CCTGGAGGTG
11151	CTTGGCACAG	TTGCGCCCCG	TCCCCATGTG	GCCGACACTC	TCAGACCAGG
11201	GCTCTGCGTG	TCCCACCTAC	GGCAGGCGAT	AGGGCTTCTT	GAGGTCTGGA
11251	GCAGGGCCTG	CATCTCAGGA	CTGTCATCCT	TGGCCTCTCT	GGCTGTCTCT
11301	CACCCACCTT	CCCTCACGTG	GCCCCCAGTG	CTTCTGTCTG	AGCAGACCCT
11351	CCCTCCTCTG	CTCCCCCTCT	TGCTCTGGCC	ATCAGCTCCC	ATCACATTGG
11401	CATCATCACT	CTGGGGCCAG	GGAAGGGGCT	GGCTCTCTGG	GGTGGTGGGA
11451	GGGATGGGGC	CAGCAGCCAA	GCCATTTCCA	GGACTTCCAA	AACAGCGCCA
11501	CTACACCCAA	CACGGCCCTC	CAGCCAGCTG	CCCACCTAGG	CCTGGGCTCC
11551	TTACAGAGCC	CCCAGAGTGC	CTCTGTGGGG	ACCCCCACTT	TCCTTCTGGC
11601	CAGTGACACC	ACCCAGCCCA	TCATCAGAAG	ACATCTTTCT	CCATGGCAGG
11651	GACCAGGGGG	TCCAAGGGGC	ACCCATGGTG	CTAGGCACCA	GGGCCTGGGC
11701	ATTCTTCCCA	TCTGGCAGCT	GGGGATGGGT	GCCCCCTGGA	CCCGTGTGTG
11751	TCTGGGGTGG	GTCACTGCTT	CTGCAGGACT	CCTAAACAAC	CTTCTGGGCT
11801	GTGGTGAACT	CTGAGCCTGC	ACCTAAAAGA	CCTGTAGTTC	TGGTCTAGGG
11851	CCTCCAAGCA	GTGTCCAGGC	AGTGTCCAGA	CCAGGGGGCG	GTCCCCAGG
11901	GACCTTGTAA	GATGTTTCTT	CTGAGGACGA	GAGCAGGCCT	CCTGGGGACC
11951	TGGGGGATGG	TCTTTTGAAG	GGCAGCAGCC	CTGGAGCAGG	GTGGGAGAGT
12001	CTGGGGCCAC	CTCTGCCCTC	TAAGGCCACC	TGAGAGGTGA	GGCCGGGGCC
12051	TGACTGGACG	TCCAGTCCCA	GAGGGGCAGG	TGCCCTGAGG	GAATGTGGGG
12101	GACAGGAATG	CTCTGCCTGG	GGCCAGGCCA	AGGTTCTCTG	AGCCCTGTGC
12151	GGATGCTGAC	AGTCTCTGGG	AACGCCTCAC	CCTGTATTTT	GGATGACACC
12201	GGCTGCTGCT	TCAATTGGAAC	CAGCCAGTCC	CATTGTTGTT	TACGTCTTGG
12251	AATTTCAAAA	AGCCCATTTT	CCTCTCTTGT	TAAAGAGTCA	GCTGAGCATA
12301	CCAGTCTCTC	TGCCAGGCTC	ATCTTGCTGG	GAGAAGTGGA	GCCCTCATGT
12351	GTTGGGGATG	CAGGTTGGCC	ACAGCACTAG	GGTGGCAGGG	CCGGCTCGG
12401	ACTCCGTGCC	AGCCTGTGCT	GGCTGCCGTG	AGAATGCCACC	CTGGTGAGGG
12451	GCGCCCTCCC	AGGGACCAAG	ACAGAACTGG	GTGTCTTCTC	CGGTCACTGC
12501	CCCATGAGGT	CCACAGAGCT	GGGGCCTGTC	AGCCGCCAGA	GGGCATGTCC
12551	CCTGAGCCCC	TGGCCTTTAA	GCCCCGTGGA	AGCAGCCGAG	GCAGAGATCA

```

12601 GCTTCAGAGC CTGGGCTGGT CCTGACACAG GCCCAGCCCT GTCCACCTGC
12651 CCTCAGCCAC GTCCACCTA TCCTTGGCCG CATCCTGACC CGCTGCCTCC
12701 CGTGTTCCT CAGGAGTCTT CTGACAGTGC CAATACCACC ATAGAGGATG
12751 AAGACGCTAA AGGTACCTGC ACTTGAGTCC TTGCCCCCCC AGCGGCCTTG
12801 GCATTGCTGG GTTGCTCTTT GAGGTGGGTG GGAAGTGGGC AGGGTCAACT
12851 CTCCTGCGAC GCCTAGTTTA TGCATGTGTT GAGGGGCTCA GGGACCCTGT
12901 AGCTGTAATC CTGCTCCAAG CCTGGGTGTC AGGCCTGCCC AGAGCGGAGA
12951 AGCATGGCAG AGATGACCGA CAGCTGGGCA GTCTCGGTCA CCGCATCCAA
13001 GTGAGGAAGC CACGGCTTTG CATGGAGGCA GGTCTCCAC ACCAGGACCC
13051 TCACGGGGAA ACAGGCCCAT GGGTAGAATT TGTTCCAAGA TGCTGTCTCT
13101 GTCTTAAAGC TCCTTAAGCT TCGTCTTCTG TCCAGCATGC ACTTGCCAAG
13151 TGGCCGGGCA GCTGGGTGAG TGTTCCTGTG TTTGCCTTTG CTTAGCCAGG
13201 AGTGTCTGCG TGCGGTGGGT TTCTGCACCA CAGATTCCAG GGGCCCCCTC
13251 CTTGCTCACC CAGGCCAATG TCTTGTGTGT TCCCCAAGAG GCCCCCAGGG
13301 CACCAGGCAC TGGGGCATGC TCCATGGATT CTGCCGCCTC CAGACCACCC
13351 ACATGGGGCC TCCTGACCCT CATCGCTCAC ACGGTCACCT AATAAGCCTT
13401 ATGCTGTTCT CAGGGCTACC CTGGTGCCCA AAAAGGGTCA GCCACTCTGC
13451 CAGTTTAGGG GAGAAAACTT CTCACCTGTC CAAAGCATAG CCTTGCTCCT
13501 GCGCGGCTA CCCAGCTATG AACTGTCCC TGAGCAGAGA TGAGCACAGG
13551 ACTTTGGGCC CTGGATGCCG GAGAGTGGGT GTTGTGTGTA TTCCCTGCA
13601 GTCTGGAACA GGCCCCAAAG GCAACAGCAT GAAGGCTGTC CAGAGGTTCT
13651 CCATCACCTC CAGCCGAGTG GGTGCTGAG CAGTGAGGGA GGGGACCTGG
13701 GAGGGGGGCC CAGCTGGAT CCTGCAGGGG AGAAGAGAAG ACAGCCAGAA
13751 GCCAGCAGCT GTGGCTCAGA TCTGAGCCCG AGCAGCCTCT CGAGGTGGAG
13801 GCAGACACCC CCCACCCAC CCCGTGCAGA AAGAAGCCTT GCCAGCCTGC
13851 CCTGAGGCTG GTACAGAGTC CAGGCAGGCT CAGTGGCCAT CATGCCCTA
13901 CGATGACTGT CACTCCCTCT CCGTGCCTCT GGCTCTGCTT GGCTCTGGCC
13951 AGGGGTGGTG ACAGCACTAG GGTGGCAGGG TGGCCTCTGA CTCTGCGCCA
14001 GCCTGCATGC CCCTGTGCTG CCCTGGCCTC TGCTGGCTCT GGCTCTGGCA
14051 CCGGTCCCGT GTTGCTCCTT TCAGCCTTCA CATACTGCTT GCGGCCACCA
14101 CAGGCCCAGG ACCCCACAG GGTGGCCACC CCACCTCCAC CCCAGGAGCC
14151 CCAGGTATCC AGCTGTACC CCCTCCCTCC CTCTGGCCTT CCCCCTGTCC
14201 TTCTCCAGTT GCCTTCTTTT CCTGCGGGCG CACCACCCAC CTGCCTGCCT
14251 CACCTGTTTC GCCTCAGCCC CCAGGTCCC CGACATCCTG AGCTCAGTGA
14301 GGAGGGGCTC GGGAGCCCCA GAAGCCGAGG GGCCCTGCC CTGCCATCT
14351 CCGGCTCCCT TTAGCCCCCT GCCAGCCCCA TGTAAGTAGC CTGGGTCTTG
14401 CTGCTGTGGG GGTGCTGTTG GAGGGCTGGC AACCCTTAG AGGGGCCACT
14451 CCAGAGCCGA GGGCAGGCTG AGCGTGGACC CTGGCTCCAG CCTCATACC
14501 CCACAATCCC TCACTGGGGC TTTCCAGGGT GGCCCCAGCC CATCGAGCCC
14551 CACCTCTTTG TGAGGAGGGC CTTGGACCAC TTTCTGCTC AAGGCCACTG
14601 GGCAGGATGG GAGGCCCTGG AGGCTCGGGC CTCAATTCCA GTCTTCAGGG
14651 TCGGTGCAGG CCTCACTCCA CCTCAGCTTG CGGGCGGGGG GGCTCCCTGC
14701 TATTGAGGCA GGCTCTGATT CAGGGCCTGA TCCCAGGGCC CAAGGGGTCT
14751 AGAACACGGG ACCCTCCCA CTGGCCTCCT CCGCCTTGCC GCCGCTCGT
14801 GTGTCTGTCT GCCTCATGTT CACGTCTCAT CTGTTCCACC CCAGCCCCCA
14851 GGGATCTCTG ACATCCTGAA CTCTGTGAGA AGGGGTTTCA GAACCCAGA
14901 AGCCGAGGGC CCCCTCTCAG CGGGGCCCCC GCCCTGCTG TCTCCGGCTC
14951 TCCTAGGCCC CCTGTCTTCC CCGTGTAAAGT AGTGGCCCCC AGGCCTGCCG
15001 CCTCTGTGTC CGGACAGCTC CCTGCGAATG GCGGCGCTC AGCAGCTTCC
15051 CACCTGCATG CACGGCCCAG CTACCCTGCC CCGGCGCCGC AGCCTGGAGT
15101 CCTGCCCTGG CGGGGCTTCC TGTGGGCTCC CATGCTAACC AGCAGGGCAG
15151 CTCCTGGCTT CTCCCTAAGG GGCCCAGACC CCTCCACGGC TCCTGCTCCC
15201 ACTGCCACTC CCCGCTCGCT GTCCAGCCCC AGGCCCTCT CCAAATGTC
15251 TGTCCCAGCC CTGGGCAGCC CTGGCCCCCT CGAGGCCCCC CATGCCCTA
15301 GGCCCTCTCT GCTGATCACT GTCCCAGCCC CACAGACTTC ACACCCACCC
15351 AGGGGCCCTG CCCATGGTGC CCAGGAGCTG CACTCAGGGC CACCCTGGTT
15401 CCTGATGTGG CCCCACCCCT TGAGCACCCCT CCCTCAGTCT AGGAGGCTGA
15451 GGAAGGTGCC AAAACTGGAA CCCCAGACCAG GGTCTCTGGA GCTCACCAC
15501 AAGGGGATAG TACGGAGAAT CATAAGCCTG GCCTCTGCTG ACCTGGGCTG
15551 TCCTCATGGG GCCAGGCCAG GCCTCCTCTG TAACGCCCGT GACTCCCTCC
15601 TCTCCCTGTA ACCCCGTCCA GCGTTCTCTA AGGGCCACTT ACCTGACAGC
15651 TTCTTGCTGG CCAGCAGCCT CTCCCTGGAG GGTGCCCTCT GCCCCAGCA
15701 GCTTCAGCCC ACGCCACCCG ACAGCCAGAG CATCTGCCCT TCACTCCTGC

```

FIGURE 3, page 5 of 16

AGCCTCCTCT	CCACGCACCA	CGCTGTCCGC	AGCAGCACCC	TCTGTCCCCC
TGTCTCCCTC	CGTCCCCCCA	TATCCCCCTC	GGTCAGCCTA	CAACCTCTCC
ACGTCCCCCT	AAGTCCACGC	TCTATCCCTA	CATCCCCCTC	TGTCCCCCAA
ATTCCCCCTT	TTCCCTCATT	TCCATTTTCC	TCCCCAAACT	CTGCTCTGCC
CCTCACATTC	TCCCTCTGTC	CCCCACACCC	TCCTCTGTCC	CCCACACCCT
CCTGTGTCCC	CCACACCCTC	CTCTGTCCCC	CATATACCCC	TCTGTCCCCC
ACACCACCTT	TGGTCCCTTG	CACGCCCTTT	TCTGTCCCCC	ACACCCTCTC
TGTTCCCTAC	ACTCTCCCTC	TGTCTCCAG	ACCCTCCTCT	GTCCCCCACA
CTCCCTCTGT	CCCCCACACC	CCCTGTCCCC	CACACTCTCC	CTCTGCCCCC
CAGACCCTCC	TCTGTCCCTT	ACACTCCCTC	TGTCCCCCAT	ATCCCCCTCT
GTCCCCCACA	CCCTCTCTGT	TCCTCCACCC	CCTGCCCCCC	ATACCCCTTT
CTGTCCCCCA	CACCTCCTCT	GTCTTCCACA	CCCCCTCCTG	TCCCCCACAC
CCCCTCTGTC	CCCCAGACTC	TCCCTCTGTC	CCCCACACTC	CGTCTGTCCC
CCACACCTCC	TGTCTTCCAC	ACCCCTTCTT	GTCCCCCACA	CCCCCTCTGT
CCCCCATACT	CTCCTCTGTC	CCCCACCTCC	CCTCTGTTC	CCACACCGCT
TCTGTCCCCC	ACACCCCTCT	TGTCTTCCAC	TTCCCTCTGT	TCCCCCACAT
CCCCTCTGTC	CCCTTGACCC	CTCCTCTGTC	CCCTGCACCC	TCCTCTGTCC
CATGCACCTC	TCTCTGTCCC	CCACATCCCC	CTCTGTCTCT	CACACTCCCT
CTGTCCCCCA	CATCCACCTT	GGTCCCCCTA	CGCACCCCCA	TCCCCCATGA
CCCCTTCTGT	CCCCCACACC	CCCTCTGTCT	TCCACACCCC	CCTCTGTCCC
CCACACCCAC	CTTGGTCCCC	TCATGCCCCC	CATCCCCCTAC	ACCCCACTT
TGTCCCCCCA	CATGCCCTCT	TGTCCCCCAC	GTTCCTTCTT	GTCTCCCACG
TCTCTCTCAT	TTCCCGTTTC	CCTCTCTGTC	CCCCAAGCTC	CCCTCCATCC
CCCACATCCC	CTTCTTTCCC	CTATATCCCC	TCTGTGCGCC	CAGGTCCACC
ATCTTCCCCC	CACACCCCCC	CATTCTCCCT	TCCTCCCCCT	TGTCCCCTTG
TGCCCCATCC	CCCACATCTG	CCTCTGTGCC	CCTCAATCTC	TGGCTTGGCT
GTCTGCCCCA	GGTTTCTCTC	CTGCGTGCCC	CCCGTGCCTG	CCTTGTGTTC
ACGTCTCGTC	TGTTCCGCCC	CAGCCCCCAG	GATCTCTGAC	ATCCTGAACT
CTGTGAGGAG	GGGCTCAGGG	ACCCAGAAG	CCGAGGGCCC	CTCGCCAGTG
GGGCCCCCGC	CCTGCCCCATC	TCCGACTATC	CCTGGCCCCC	TGCCACCCCC
ATGTAAGTAG	CACCTTGAGT	GGCCGTGGCA	GCGGCTGCCT	GGAGGGGCTC
GGGGCGTGCG	AGCCTGGCAG	TGGTGCTCTG	GGAAGGGCCA	TTCTTGCGGA
GGAGGGCGGG	GCACAGGATC	CCTCTGTCTG	GTCCCAGGGA	ATTGCTTTGA
AGCACATGAA	GGTGCCACTG	GGTCTCAGAA	AATGGAGGTT	ATGGTTATGA
AGTGTGTATG	ACATATGTGT	ATAGGAAGAG	CGTCCGAAAG	AGCAGGTTTG
TTGCCGACCC	CAGCATTCGC	AACCCTGAGG	TCCACAGCTT	TCTCCTGATG
GGAGGGGAAT	GGGTGGCAAA	GGGTCTGCGC	GTGTGGCAAG	GGCTAGCACG
CCAGGAGCTG	CTGGCTTGGG	TCAAGGTGGA	CCTGCTGGGC	CGGGACAGAA
AAGTGTCACT	CCCGGCCCTGA	GACGCTCTAG	CATTAGAGCT	GTCCAAGTCC
AGACAGCAGG	GAGCAGGTGG	GGATCGGGAG	GCGCGGATCT	GGGGGGCAGC
TGGGGCCAGG	CTGAAACAGA	GCGGGCGGGA	CAGGAAGCAC	AGGCTGGGCA
GCCTCCCCGG	CCAGGGAGGA	GCCAGGCTGG	GCCACCTCCC	GGTCTGTCTG
CCGACTACCC	GCAGTATCAC	TTACAGGGAT	GGATGACATC	CCAGGGCTGC
TGCCACCCCC	ACCTGTGGGG	AGACACCAGA	CTGGGGGTGG	TGTGGAGATA
CTCTTAGAGA	AGAGGCTGCT	GGGCCACGGG	CTCGGCATGG	CAGGGCAGTG
GCTAGGTAAG	TACTTGAGGG	ACAGGTGGGG	TCTGCTTGCC	ACCGTCCCCCT
CTGCAGGCTG	GGCCTGGGGG	CTGCTGCAGG	CGGCCAGGGC	AGAAGGGTGT
GGGGAGAGTG	AACCCACAGG	AGCAGCGGCT	CGAGGAGGGG	GATGCAGGCT
GCAGGCTCAA	AGGGGCACTG	GATCCACCCT	GGGTGCCCCG	GAGAGCAGGG
GGCAGCCCCC	GGAGGGGTAC	TCACCCCCAG	AGCTTCTGTG	GTGCGCTGAG
GACCCCCAGC	AGGGGTTGAC	TGAGGGGATC	AGAGGCAAGC	AGCTGAGGGG
AGAGGCCAGG	TTCTTGATGC	TGATAGGGTC	GGGGTGCCTG	GGCGACCAGA
ACTCAAGGAG	GGAGGCATGG	GGAGGGGGCC	CCGTGCAGCT	GGGGTGGGTG
CACCGCAGAG	CCTCTGGGAG	TGGTCAGAAC	CCCCGACACC	TGCCACTTCT
ACAGCAGCTC	ATCTGATTTT	AAGGGGCTTG	CTGCCCTTGC	AGAAGTGGAG
GGGTGTGCCC	AAAGGAGCCT	GCCTGGAAGG	TCACCCCATC	AGGTGTGGCAT
GACCCACAGC	CAGGACTGCA	GCCTGCCCTC	AAGGTCTGTG	CAGTATCTGG
GGTGAGTCTT	CTGAGGACAG	GGCCCAGGGT	GGGTGTGGAG	TGGCCAGCTC
GGGGTCTCGT	GTCCAGGCTC	ACCTTCAGGG	GCCACAGCAC	AGACCTGCCC
TTCCAGAGTC	TTCCTTGAGC	TTGGCTGGGG	AGGAGGGGGC	TGCAGGAAGG
AGCTGTGAGC	AGGGCAGGAT	GGAGATTCTG	GTGGCCCTCC	TGGGAGGGGC
TGGGCAGGGC	TGGGAAAGGG	GTGGGTGAGA	TGTTCCGGAA	CTCAGGGAAA
GGAAGAGTCT	GGGTACTGCC	CTGGGGGCAC	CTGGGCCCAG	GTGGCAGGTG

FIGURE 3, page 6 of 16

18901	GCCAGCTTTC	TGCTCTCTTT	CCACCTCCTT	TCTCCAGAAG	GCACCCACCA
18951	GCTGTGTAAA	TAGGGCAGGT	GCCCACGGCC	CGCCTCAGGC	CCCGTCTCCT
19001	CCCCACCCAC	GCTCTCTAAT	CGCGGATTAT	ACACAATCCA	GCCTGATCCC
19051	TGGGCAGCTG	CCCTCCCTCC	CGCAGCCACC	TCTGGCTCTG	AGAGATGGGC
19101	TTGGGGCCAG	CCTGGGGTCC	CAGGAGTCCA	GGCCAGGATG	AGAACCTGCT
19151	CTGACCCCA	CTGGACGCAT	TAGGCTGCCC	TGGACCTGTT	GCCTCACCCC
19201	AAGAGGCCA	CAGGCAATGC	AAAGGCTCCT	GTTCATGTCA	GGGCACCTGG
19251	AAGGCTGAC	TTGCAGAGGC	TCTTGGCTCG	TGCAGACCCC	TCCAAGCCCA
19301	GGCCCTGCC	ACCACCTCCC	CTTTGTCTCT	GGAACTGCCA	GGACAGCTTG
19351	TCCTCAGCCA	GCAGGTTTCC	CGACCCGGGC	ACCTCTTCAT	GT'TGGGCCCC
19401	CCTCCTTTCC	CTCCATCAGG	GATCATGCCC	TTCTTCAGGG	GCCTGGATAT
19451	CAAGGACACA	AAAGCTCCCA	TGTGCTATGT	GGGGAGGCAG	AGTGGGGGCT
19501	GGGT'TGAGCT	GGGTGTGTTG	GAGCGCATT	CCCGCAGGCA	GGGGCAGCCT
19551	AGGCT'TCCCA	TCTGTGGAAT	GGGTGGGTGG	GTCTCACAAC	GGACCTGCTT
19601	CCCGTACTTC	AGCACGGTTA	CCACTCTTGA	TTGGAAC'TCT	GACCATGCAT
19651	CTCCTCTTCT	GTTTACTTCA	CGCTT'TCTCT	TCCCATCAAC	TCCCATT'TTA
19701	ATTACAATTT	GTTTAAAGC	ACTGCATATT	ACTTCATTAA	ACAGAAGATT
19751	AGTTTCACTT	ACCATTAGTG	TAAGGTGACT	ATATAACCAA	AGCAGACTGG
19801	AAACCAATG	ACATAATGTC	ATTCTCTTCT	CCATTCCAGC	TGCCTGCTGC
19851	TGTGCGCTG	AGAACCCTG	TGGAGTGGGA	GGGGCAGCTG	TCTCTGTACA
19901	TTAGAAAGGG	AGGTTAACTA	AGTGACAGGA	GGTGT'TTGGG	ACATGTGGAC
19951	ACCAGACTTC	TCTCTTGATG	CAAGGAGGGC	AGAGCCAGGC	AGCCTAGTGG
20001	GGGCTGGCTT	GGGGGCTGCT	GGAAGGACTG	GCTACAGGTG	GAAGAGAGGT
20051	CAGACCTGAA	GCTTGGGGCC	ACCTCCAGGA	AAGGACAGGT	GAAAGTGGAG
20101	GCATGAGGCA	GGGGAGGGC	AGGTGCCAGG	CAGAGGTGG	AGAGAGGCA
20151	GGAAACATAGC	AGCTGGGGCG	GGGGCGGGCC	CTCAAGTGTG	ATATGCTACT
20201	TTCCTGGGGC	CCAGGGGCAA	GGACAGGAAC	AGCCACAGCA	TGTGTTGGGA
20251	CAGAGCCCTG	TGCCTTCCTA	GAGCTGGGCA	GGTGGAATGG	GGCAGGAATG
20301	GGACTCGTGG	TGGCTGCAGC	AGGAAC'TGGA	GGGGAAGGGG	CTTCTGGATC
20351	CTGCAGCCTA	CCTTCTCTAGA	GGCCAGCTTT	CCGGGGTCCA	CCAGTGGGTT
20401	GGGAAGCTGGG	CTTGTGTAGC	AAAGCTGCC	TGAGGACCAT	GCATGACATG
20451	GCTAGATGA	AAGTTAGGAA	AGAAAGGGAG	ACAAGCTGGC	AGCAGAAGTA
20501	CAGCTGGGTC	AGGAGCAAGG	GCCTT'TCCAG	ATAGGGACAA	CCCAAGAGTG
20551	CACATGTGCC	CACGCCACAC	AACACAGGCA	CACACGACAC	GTGCACGCCT
20601	ATAGGCACTG	CACACACACA	TGCACAGGTG	CTCATGCATA	TGTATGAGCT
20651	TCATCTACAC	ACATTACAT	GCCGTCTGTG	TTATGTGCAT	GTTTCCATAC
20701	ATGCACATGA	ATGCACAATC	ACGTGTACAC	ACATGCATGT	GATCACATAC
20751	ATGAACATGT	GTGCACCCCA	CTCTCAGGT	GCATCGGGC	TCCTCCTGCT
20801	GTCAGTGTGC	AGCAGGGGAC	ATGAGGCCCC	AGAGCAGACA	GGTGCAGCAC
20851	AGGCGTTCCC	AGGCAGTGCC	CCACACACAT	GCATGAGCAC	ACCCGGGCAT
20901	GTGGCGCCTC	CTTGTGTGGAC	TCAGTCCACC	TGCCAGGTGG	GCTCCCTGGT
20951	GGTGTGAGCT	CCCAGAGGTC	TGGCGAGGTA	GATAAAGGCA	ACCCACCAC
21001	CAGGCGTGCT	GAGAATTCCC	GTCTCTGGAG	GGGCACAGTG	GCTCATACCT
21051	GTAATCCCAG	CAC'TTGGGA	GCCCGAGGTG	GGCAGATCAC	TTGAGGTTAG
21101	GAGTTTGAGA	CCAGCCTGGC	CAATATGGTG	AAACCTCATC	TCCACTAAAA
21151	ATATACACAC	ACAAAAATTA	GCTGGGTGTG	GTGGTGTGCA	CCTGTAGTTC
21201	CAGTACTCG	GGAGGCTGAG	GCAGGAGAAT	CGCTTGAACC	TGGGAGTCAG
21251	AGACTGCAGT	GAGCCGAGAT	CATGTCACTG	CAC'TCCAGCC	CGGGTGACAG
21301	AGTGAGACTC	CATCTAAAAA	AAAAAAAGAA	TTCCCTCCTC	TGGGAATTTA
21351	GACCACAGAC	AGGTGTCATG	TATGTGGCCG	TTGGAGGCAG	CACTCACAGC
21401	AAAGAGTGGA	AACGTCACCA	CAGGGCCTGC	CTTCTGGTGA	AAATGGTGTG
21451	CTGCAGGGCG	GGCAGCTGTT	TGAGGGCAGG	TG'TCCAGGT	GCGGCC'TGCA
21501	GCAGCTGAG	GGTCACAGAG	CGCAGTGTCT	GGAGTGCAGA	GACTTCCCCC
21551	ACAGGGGAGAG	TTCCCGAGAA	CCTGCTTCCG	GTGCACTTCT	GGGGGT'TTGA
21601	GTTTTTTTCCA	CGGACGAATT	ACTTTTGAGAA	ACCAC'TGTTA	CTCTGTGTGA
21651	TAGGTGAGCG	TGCGTGTGCA	TGTGTGTTCT	GTGTGTGAGT	GTGCATGTAT
21701	GTGCGTGCCT	GCGTATATAT	CCTCGCAGAT	ACGGCTAGGG	ACCTCACTCA
21751	GGACAGTAGT	TCTGCCTGAG	GAGAGTGAAT	GCGGCAAGAT	TGAGGAGAAC
21801	ACAGGCATCT	TCAAAC'TACA	TGTGCGGTGC	TTTATT'TCTT	TAAAAATGCG
21851	TCTAAAGCAA	ATAGGAAAAA	GTTAAGATTT	GAATCCGTAG	AGTGTGGGTT
21901	CTATTATTCT	CTCCACATCT	TCCATACGTT	TAAAACTTTT	TGCAATGAAA
21951	ATAAGTGTGA	GTTAAAGCAG	CAATGCAGGC	TGCCAGTGAG	CGCCCCGGAG
22001	GCCAGTGGAG	ACCAGCATGG	CTGGGTGGCC	TGTTGGAATC	CAAGGGGGGC

```

22051 GGGCAGGAGC TGCAGGCAGG CGCCCGGGAG TAGCCCGGGC ATGGGGGTGC
22101 GGGGCAACAG GGATGTCTGC AGGGGTAGCA TGTGGGCCCC GGAAGCAAGA
22151 CAGGTGGAGC CAGCCGGATG CGGCTCCTAT GAGAAAAGCG GGAACAAGA
22201 GACCACGCTC GTTCTTCCTG CTGCGGGGAC AGCCCTGGTC ATCGCTCCGG
22251 GGAACCTGTC AGCCTGCGCC GCACGTGGCC GCCCCCTGCT GCTTCCTCCT
22301 CCCC GGCTC CGGGTGGCCT TGCTGACGGC TCCTTCTCTG AGGCAGGTCT
22351 CTGCCTTCTC GCCTGGTGCC TGCACTCAGT AGCCCCCTCA CCAGAGCTGC
22401 TGGGTGAAGG AAGCACTAAG AACCAAGGC TCGGGAGGAG AGTGGGGCCG
22451 GGAAGCTGCA GGAAGCGCA GGGCCAGGCC TGGTGGGCCC AGGGGCTGGC
22501 TCACGGGAGG GCAGGAGGGA GACTGTGGCG GACAGCACGT GGGGCCAGGA
22551 GGTGACCTCC AAGTGGATTG TGGGTGGGTT TTTTGTCTC TTTCTGCATT
22601 TTCCAGGCAT TTTGTAATGT GGATAGAATA TTTCTGTTCT TCAAAAATAC
22651 TTTAGTTAAG AAAAATAAGA TGGAAGCTGT TGCACTTGAA AATGAGGAAG
22701 CCACTGGTGA TGCAGGGGGG GCGGCGGAGA GGACCTCTTC TGCAAATAGC
22751 GGCAGGAACA CGGCATGGAT GCAGCTCGCG CTCCCCCAGG CCCTCCCCTG
22801 GGCTGTGTGG AGGGGTCCGG GGGGAATGGG CCAGCGCCCA GTGGTCACCT
22851 GGCCATGTCT CCCACAGCC CGGAAGCAGG AGATCATTAA GACCACGGAG
22901 CAGCTCATCG AGGCCGTCAA CAACGGTGAC TTTGAGGCCT ACGCGTGAGT
22951 CCCTGGGGCT GGGGGGGGGC TGTGCAGGAC AAGGATGTGG GACCCTTGGG
23001 GGGGCTGCT CAGAGTCAGG GGTCCACGGG GCCCCCTCTC ACTTGATT
23051 GGCCCCCAGG AAAATCTGTG ACCCAGGGCT GACCTCGTTT GAGCCTGAAG
23101 CACTGGGCAA CCTGGTTGAA GGGATGGACT TCCACAGATT CTACTTCGAG
23151 AACCGTGAGT GAGGAAGCCC GGTGGGCAT GAGGGGGCGG TGCCCCCAGG
23201 AGAGCCTCTC GGCCCTCCC AGGGACAGCA TGGTGGCTGC CTATGGAAGC
23251 CTTGTCCCCT CTGTGCCAG GGTGGCCAG CCACCTCTCC CCCGCCAGAG
23301 GCCATACCCA GCCCCAGAA TCCACTCTT GGAGGGGCCC ATGCTGCTCC
23351 CAGGAGAGCC GAGCCTCCCC AATAAGGGGA GTTGAAGAGG GGAAGGATT
23401 AGGCTGGTGG GGTGGAAGAC GGGCACCAGG GCAGTCATGG TAACCCGAGA
23451 CCCCCGCCCC GCCTGCTGTC CACAGTGCTG GCCAAGAACA GCAAGCCGAT
23501 CCACACGACC ATCCTGAACC CACACGTGCA CGTCATTGGA GAGGATGCCG
23551 CCTGCATCGC TTACATCCGG CTCACGCAGT ACATTGACGG GCAGGGCCGG
23601 CCCC GACCA GCCAGTCTGA GGAGACCCG GTGTGGCACC GCCGCGACGG
23651 CAAGTGGCAG AACGTGCACT TCCACTGCTC GGGCGCGCCT GTGGCCCCGC
23701 TGCAGTGAAG GTGAGTGTTT TGTGCTAAGT GACAGCTGGG GCAGAGGGGT
23751 GGGGTTGGTG TGAGTGGCTG CAGCCTGGGG AGGCGATGGG GAGCGGTGGG
23801 GCCTGTGGCA GAGCCATGC CTGGGAAGTC CCTGAGCTTT CCTGGTGAGG
23851 CCACAGGAAT GATGTCAAAT TAGGGACCAC GGCAGGCTGG GTGTGGCAGG
23901 CCTCCCCAGA GGAAGGGGA GCTGGTGAGG GCCTGAGCAG TCCACTCTGG
23951 CCAGAGCTGG GTGGGTTGCA GGTGGATGGG CCCC GGGCAG CACAGTCTCTG
24001 GGCACCATGC CCTGTTTGTG AGGACTGTTA GAGCCCCAGA TGGGCGTTCC
24051 CCAGGTGGTG GGTGCAGCGG GCCCAGAGCC CAGTTTTTACA GGGATAGTAG
24101 TAATTGGGTT GGGCACCTTG AACCTCTCTC CCGAGTGGGC CCTTTTCTGG
24151 ACTTTAACC TCTCTGCAGT GCCGCATGGC AGACAGCAGA GCCTGGGGGT
24201 GGATGGGAGA GGGGGCTGCT GAGGAGCTGA CCCACCCGCC CCATTTTACA
24251 GCTGCGCCCT GGTTCGCGG GACAGAGTTG GTGTTTGGAG CCCGACTGCC
24301 CTCGGGCACA CGGCCTGCCT GTCGCATGTT TGTGCTGCTC TCCTTCCCTC
24351 CCCTGGTGCC TGTGCTGCA GAAAAACAAG ACCAGATGTG ATTTGTTAAA
24401 AAAAAAAAAA AAAAAAAAAA AAAAAACAAG ATGACGACGA CAACCACAAA
24451 AAAAAATTGAC ATCAGATGAA ATGAAAAAAA AAAAAACAA AAAAAACTAA
24501 AGGAAGGAAA AAGCTGTAAA AATCACTGGC ATTCGTGGGG CCACTCCCCA
24551 CCAAGCTCC ACGTGTGTCC GTCTGTGCTC CTGGCCTCTG GGGGACCAGC
24601 TGGGACATGA ACTTGTCTGC CAGGCCCCCG TCGCGTGTG AACGGTGTTA
24651 GTTTGTAGGT AACGCACACA CCCCACACCT AAGGTGTCTG CATCCTCCTG
24701 CCAACGCATG GGCTCCACGT GGTGTGCTCG CTGGCTGTG TGAAGTCTG
24751 CTGTCTCTTG GGAGGGGCTG TGGGGGCCCC CTGGGCTGCC TCCTTTCCCG
24801 CTAGTTGTGC CTGAGAGTTG CTGTTGTTCC TGCTTTCCCT TCCCTTCTCT
24851 TCATCCCCTG AAGGGCTAGG TGTGGGTTTT CCGTGCCCGG TATCCCCACA
24901 CACCCAGCAC GGACAACCCT TCGGCAGAGC CCAGGCCGGC CCCTACCCCC
24951 CTGGAGTATT GAAACTGGAG TCCCGTCCCC AAGGCCTTCA GAGATGCCCC
25001 TACACACCCA GGGCTCCAGC TCTGGTCTCT CTGGGGGAGT AAAGTGCAAA
25051 GAGGGGCACA GCTTAGTTTT GGGCCTCTCG CCGAGCAAGA GACAGCACTG
25101 CTGGCTACAG CTCCAACACA GCCAGCTGTG GCAAGAGGAC TCTGCCTGGG
25151 CTGGCCCCC TCCTGTGTGA GGTGTCTGTC CCTTCTCTGC TGGCCAGCAG

```

FIGURE 3, page 8 of 16

25201 CAGATGCACT GGCAGCTCCC AACCCTGTTT CCGCCCCTCG GCCCTCCCC  
 25251 AGCCTGTTTC GCTTCTCTGC AGCCCGCAAG GGGGAGCAGA CTTTTGACAA  
 25301 AGGACTGCGG GCCTCGCTCA AGTCCCTGAG CCCCAGCTG AAGCTGGGAG  
 25351 GGGAGGCCAG GCTTTGTGTC TGGGCATATT CGTCTGCTGA TGGGGTTTGG  
 25401 GGAAGCCTGG GGCTTGGGGT TTGGTCGGGT GGTGCAGCTA GTGGCAGAGC  
 25451 GGGATCAGAG GTGGTGGCTG CCCAGCTTCT GGGCTGAGAC AAGGGTCTGT  
 25501 GCAGGGGTTT ACTGAAGTGG GAGTGCCTTT GGAATCTGGG CCGGGAGCAG  
 25551 AAGGGAGCAA AAGCTACAGT GGGAGCCAGC CTAGGGCACA TGGGAGGCGT  
 25601 GAGGGCAGTG CTGCCCCTGC AGTGTCAAGT GTGCCAGTGC CTGGCGGGC  
 25651 TGCAGTGCCT GTGAGGGCAC CTCTAGGTG GGCCAGGGAT GCAGCTATGG  
 25701 AGATAAGCG GGCTGGGGAC AGAAACAGGT GGGCACAGGG CCCAGGACAC  
 25751 CAGCGGATGG AGGGCAGGGT CTAGCCCTGT GCTCCTGAGC GTCGGCTGCC  
 25801 TGGGTTTCGAG GCGGTGGGTC CCCGGCCCCT TGTGATGGTG TGTACCATGG  
 25851 GGGAGCTCGG GGACAGGGCA AGCCCGAGCA TGGTGGGGCT GCAGGGTGGG  
 25901 TCTGAAGCCA GGTGGGGTGG GGGTGGTCAC AAGCCCTGAC TGCAGAGGGT  
 25951 CAGGGGCTCC TGCCCCAGTG CCTGCCCCCT TTCAATTCAC ATTGTTTTCA  
 26001 ACAAGGATTT TCTTTATCTT CCCCTACAAA TCAAGCCAAG GGAGGGGCAC  
 26051 AGAATGGGGA ACAGGACACA GGATCCTAAA CTCCAAGGGG ACTGTCCACC  
 26101 GATGAACACT CAGAGTGGAC ACCATCTTCC GTCCACGCTG TGCCAGGAC  
 26151 AGCTGTCCCC ATCCATGAAC ACAGGGTAAA CATCTGCCGG GCTCCGCACC  
 26201 AGTGCTCCC TGGGCCATGG GACAGCGGCA GGGCTACCA CGGACAGCAC  
 26251 GTGGCCAGC AGCCGGCCAC CCTGGCGTCC TGGGGCCTCC TCCCCTCCTC  
 26301 TCCCTCTCAC CTTGTACCT CCACGGAGCT GCCTGTCTGG GATAATTTGG  
 26351 GGATTTTTTT TCTGGGGAT AATTCTTTTG CATGACCCCT AAAGAGCAAG  
 26401 CCACACCGGT CTGCTAGCTA GGTGTCCGCG GTGTGGTGGT GGCGGCCGCT  
 26451 GGCCAGCGCT GCAAGGGGTC GGCTGCCAC GGTGCTGGCT GGCTCCCCCT  
 26501 CCTCTCTCTT TTTGCTGAGT TTCATTGTCT TTTCTTTCTG AGCCTTGTA  
 26551 GTGTACAAAA ATTATTCTTA TTTTGTCTG TCTCGGAAA CTGCAAATAA  
 26601 AAGAAAAACA GGACAACTG CTTCAAGTGC AGCTGGGTGC TTTAGCTGGA  
 26651 ATCCTGCCGA CCTCTGCGC CAAAATACAG ACTCAAGCCC GGTCCCTGGC  
 26701 CAAGACCTTA CTTGGGCCCC TCCTCCAATG AAAGGTAGTG CTATGGGAGC  
 26751 CCTGAGCTGG CCCTGACAGT CCTGAGCCCC TCTAGGGTGA ACGGCTCACC  
 26801 CCAGGTAGGG CACTAGTCAT AGATCATAGC TCTACCAGCT GTCTCCACCT  
 26851 CTTCTCTGCG TCCTCTGAAG TCTTCTGGG CCAGCGCTGT CCACCCTGAA  
 26901 TGCTGGAAGT GAACTGGAT CCCAGCCCC AACACCCCTG ACCTCTCCAT  
 26951 TCACCCCGG TGGCCGCTAA GGATGTGGCC AGGGCAGCCT CTGGGCAGGA  
 27001 AGGAGCCCCA GGACCAAGAC CTCTGGCTGT CTTGCTGTTT CTTCCGCCCC  
 27051 CTGCTACATG TATTGGCTAT TCTGGATGCT GAGGACACAC AGTGACCACA  
 27101 GAGCCGGGCT CCACCCAGT GGATTATGCA GACAGATGGC ACGCAGGCCT  
 27151 GTGTGGACAT CAGCCTCGGG CACCAGACAT AGGCAAGGCG CAAGGTGATA  
 27201 CAGTAGGCAG CCACCATGGG GGCCAGGAGG CTCCAGCAGA GGCCACACAA  
 27251 CCAGCCCGA ATCCAGGACA GAGAGCTGGA ATGGAGACAG GGAAGCCAGA  
 27301 TACCAGGCCA GACTGGCCAG GTGCTACAGG CCTGTGGGCC AGGCCAGGCT  
 27351 TGGGGACTTC GTCCTGGGTG TGAAGGAGAC AGGCACCCCT GAGGCCTTCC  
 27401 CTCTGCATCT CCAGCCCAAG CTAAGCGCAA ACTCTTAGGT TGGAGTAAGG  
 27451 AGTAACCCCC TGCCAAGTTT CTCCTGTCTT CAGGCTCCAC CCACCACCTA  
 27501 TGCTGCCTGG CCCCATGGGG CACACGCTCA GGCCCAGCCT GGGAAAGCAA  
 27551 CTGCACCTGC CTGTGCTATG CTGGCCCTTC TCAGCCTCAA TGCCCTCCTC  
 27601 CCTCCCCGAC GCACCCCTCGT GGCCCCCGCT GGGCCCCCTG ATGCACCCCTC  
 27651 ATGTCTCCAT GGCAACCTGC TCAGAGTGTG GCCCTGCCCT TGGCTCCCCT  
 27701 CCACACCTGT GTCCCAGGCA GTGCCACGGC ACTTTCCTAA ACAGAAGGAT  
 27751 GGGCTTCAAA ACAGTCCCAG ACACTAAACA CACCTGCATT TTGGGTCCAA  
 27801 GTAACCTCTG ACAAGACGAG TGCCCCCTACA CACCCTCAGT CCTATCCACT  
 27851 ATGGGCAAGG AGCCTGAAGG ATCCCCCAGA ACTGGCTAAA GCCCTCAGTC  
 27901 TCCTCCTCCA CCCTGAGCAC CTTACGCGG CAGAGTGGCC CTGGATGTCA  
 27951 GCTTCTTGCT CCCCATGGTC TGCACCTGGA CAGGTGCTCT CAGGTGTGTG  
 28001 GGTGGGCAGG TGGCAGGTCC CAAGAGCCAG GTGCAAAGAA TCTAGGCCAG  
 28051 TGCCACAGAG TGCTGCAGTG TCTGTCCCCA GCATGGTATC TAGGGCTCCA  
 28101 CTTGCCTATC AGCTGTAATC GGAGGAGGCT TTCCAGGCCA GGCCTCCCCC  
 28151 AGGAAGGCTG CAGGCACTGC GGATCGTGCG CCCTCACATG CATTATTCCT  
 28201 GAGGCCCTTC TGCAGATGCC ATCAGGGCAG CAACTCTGAT GAGGTATTAG  
 28251 GGCACAGCAC ACAGGGCTAA GCCACCCTGT ACTGGGCCAA GCGCTACAGG  
 28301 CAAAAAGGAC ACCACCGACG GGCATTTTCAT TCATCGCTTT TATTTTTATA

FIGURE 3, page 9 of 16

28351 TATTTTGTGAG AGGGAGCCTC ACTCTGTCGC CCAGGCTGGA GTGCAGTGGC  
 28401 GCGATCTTGG CTCACTGCAA CTTCTCCCTC CTGGGTTC (SEQ ID NO:3)

#### FEATURES:

Exon: 232-340  
 Intron: 341-431  
 Exon: 432-515  
 Intron: 516-1110  
 Exon: 1111-1205  
 Intron: 1206-1424  
 Exon: 1425-1547  
 Intron: 1548-1981  
 Exon: 1982-2065  
 Intron: 2066-3015  
 Exon: 3016-3058  
 Intron: 3059-4102  
 Exon: 4103-4177  
 Intron: 4178-9088  
 Exon: 9089-9126  
 Intron: 9127-9303  
 Exon: 9304-9375  
 Intron: 9376-10898  
 Exon: 10899-10943  
 Intron: 10944-12713  
 Exon: 12714-12762  
 Intron: 12763-17130  
 Exon: 17131-17133  
 Intron: 17134-22868  
 Exon: 22869-22944  
 Intron: 22945-23137  
 Exon: 23138-23154  
 Intron: 23155-23475  
 Exon: 23476-23705  
 Stop: 23706

#### CHROMOSOME MAP POSITION:

Chromosome 7

#### ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
487	T	C	Exon	55	H	H
496	T	C	Exon	58	L	L
1662	T	C	Intron			
1785	T	A	Intron			
1889	A	T	Intron			
2416	C	T	Intron			
4698	A	G	Intron			
5424	C	T	Intron			
8722	C	A	Intron			
9982	G	A	Intron			
10951	C	T	Intron			
12603	T	C	Intron			
14583	C	T	Intron			
17290	T	C	Intron			
18188	C	T	Intron			
19911	A	G	Intron			
21328	C	A G	Intron			
21391	T	C	Intron			
22588	C	T	Intron			



22965	-	G	Intron			
23498	G	A	Exon	312	R	R
23663	T	C	Exon	367	S	S
25427	A	G	Beyond ORF (3')			
27727	C	T	Beyond ORF (3')			
27834	T	C	Beyond ORF (3')			
28336	G	A	Beyond ORF (3')			

Context:

DNA

Position

487

CACCTCTGGGTTTAAACAACATGCACCCTTGTGCCGGTCACCTCCCTGCAGCCGGAGAAC  
CTGCTTCTGGCCAGCAAGTGCAAAGGGGCTGCAGTGAAGCTGGCAGACTTCGGCCTAGCT  
ATCGAGGTGCAGGGGGACCAGCAGGCATGGTTTGGTGAGTGCCAGGGGCAGGGTGTGTG  
GCTGGCAGTTGGCAGGGCAGGAGGTGATGCTGACAGCCCCTTGTGGCCTCTTCCCCTCTC  
TCTAGGTTTCGCTGGCACACCAGGCTACCTGTCCCCTGAGGTCTTCGCAAAGAGGCGTA

[T, C]

GGCAAGCCTGTGGACATCTGGGCATGTGGTGAGGCCTGGCCTGAGTTGGTGCGGGGCAGG  
GCCTCGGGTGTTCAGGACTTCCCACCTACATCCTGGAGTGTGCAGTGGCCAGCACGTCT  
TGCTCTCATCTGGGTTTATCTGTGTGACACCTGCCCTTGAGCTGCCCTGGCAGGGGTCTG  
CCCACACAGCCAAGAGCCCCCTTCCACCCAGATTAGAATTGCTCACATGAACCTGGCGC  
ACCCAGTGCTCGCCTGCGCTCAGCAGAGGTCTGGTCCAGAAGTGTGGTGGGTGGATGGG

496

GTTTAAACAACATGCACCCTTGTGCCGGTCACCTCCCTGCAGCCGGAGAACCTGCTTCTG  
GCCAGCAAGTGCAAAGGGGCTGCAGTGAAGCTGGCAGACTTCGGCCTAGCTATCGAGGTG  
CAGGGGGACCAGCAGGCATGGTTTGGTGAGTGCCAGGGGCAGGGTGTGTGGCTGGCAGT  
TGGCAGGGCAGGAGGTGATGCTGACAGCCCCTTGTGGCCTCTCCCCTCTCTAGGTTT  
CGCTGGCACACCAGGCTACCTGTCCCCTGAGGTCTTCGCAAAGAGGCGTATGGCAAGCC

[T, C]

GTGGACATCTGGGCATGTGGTGAGGCCTGGCCTGAGTTGGTGCGGGGCAGGGCCTCGGGT  
GTTTCAGGACTTCCCACCTACATCCTGGAGTGTGCAGTGGCCAGCACGTCTTGCTCTCAT  
CTGGGTTTATCTGTGTGACACCTGCCCTTGAGCTGCCCTGGCAGGGGTCTGCCACACAG  
CCAAGAGCCCCCTTCCACCCAGATTAGAATTGCTCACATGAACCTGGCGCACCCAGTG  
CTCGCCTGCGCTCAGCAGAGGTCTGGTCCAGAAGTGTGGTGGGTGGATGGGAGTGGAGAA

1662

GAATTCTTGCCCCCTGCCTGAGAGGGAGCTTCAGGCCCGGCCGGGCGCTGTTTCTTCTG  
CAGTTCCTCGTCCCCTGAGTGGGACACCGTCACTCCTGAAGCCAAAACCTCATCAACCAG  
ATGCTGACCATCAACCCTGCCAAGCGCATCACAGCCCATGAGGCCCTGAAGCACCCGTGG  
GTCTGCGTGAGTCGCCCTTGGTGCCCATGGTGGGGAGGGGGCTCCTGGTGGAGATGGCCT  
CAGACCACTCCCCTGGCAAGGACCCCAAGAGGGTCTGTTCTTGACATCCAAGAGCTCCC

[T, C]

TGGGTCCCCTGGGTGCTCCTTGTGGCCTCTGGCTTGGACATACCAGCACGTTTGTGAGG  
CCTGGGGCTTGGAAAGGCATTAGAGGGTAGAGGTGATCCCTTCCCTCCCAACTGCAGTCCCTG  
TCTGTGAGGGGCAGAGTGGACGAGGCAAGGGAGAGACGAGTCTTGAAGTCCCAGGCGGGT  
GGGGACAGACAACCCCTTGCCGCAATGGTGGCCGGTGGCTCTTGGCAAGTGGGGACCCAG  
GGTGCCACAAGCCTTGCCACCTGGCCTCTCCCCTGTGCCTCGGGCTCGGCTGCCATATG

1785

CTGACCATCAACCCTGCCAAGCGCATCACAGCCCATGAGGCCCTGAAGCACCCGTGGGTG  
TGCGTGAGTCGCCCTTGGTGCCCATGGTGGGGAGGGGGCTCCTGGTGGAGATGGCCTCAG  
ACCACTCCCCTGGCAAGGACCCCAAGAGGGTCTGTTCTTGACATCCAAGAGCTCCCTTG  
GGTCCCCTGGGTGCTCCTTGTGGCCTCTGGCTTGGGACATACCAGCACGTTTGTGAGGCC  
TGGGGCTTGGAAAGGCATTAGAGGGTAGAGGTGATCCCTTCCCTCCCAACTGCAGTCCCTGTC

[T, A]

GTGAGGGGCAGAGTGGACGAGGCAAGGGAGAGACGAGTCTTGAAGTCCCAGGCGGGTGGG  
GACAGACAACCCCTTGCCGCAATGGTGGCCGGTGGCTCTTGGCAAGTGGGGACCCAGGGT  
GCCACAAGCCTTGCCACCTGGCCTCTCCCCTGTGCCTCGGGCTCGGCTGCCATATGACC  
ACCCATTTCCACAGCAACGCTCCACGCTAGCATCCATGATGCACAGACAGGAGACTGT  
GGAGTGTCTGAAAAAGTTCAATGCCAGGAGAAAGCTCAAGGTGAGGCCCTGGCCCCTAGT

1889

GTGGAGATGGCCTCAGACCACTCCCCTGGCAAGGACCCCAAGAGGGTCTGTTCTTGACA  
TCCAAGAGTCCCTTGGGTCCCCTGGGTGCTCCTTGTGGCCTCTGGCTGGGACATACCA  
GCACGTTTGTGAGGCCTGGGGCTTGGAAAGCATTAGAGGGTAGAGGTGATCCCTTCCCTC

FIGURE 3, page 11 of 16

CAACTGCAGTCCTGTCTGTGAGGGGCAGAGTGGACGAGGCAAGGGAGAGACGAGTCTTGA  
 AGTCCCAGGCGGGTGGGGACAGACAACCCCTTGCCGCAATGGTGGCCGGTGGCTCTTGGCA  
 [A, T]  
 GTGGGGACCCCCAGGGTGCCACAAGCCTTGCCACCCTGGCCTCTCCCCTGTGCCTCGGGCT  
 CGGCTGCCATATGACCACCCATTTCCCCACAGCAACGCTCCACGGTAGCATCCATGATGC  
 ACAGACAGGAGACTGTGGAGTGTCTGAAAAAGTTCAATGCCAGGAGAAAGCTCAAGGTGA  
 GGCCCTGGCCCTAGTCCCAGGCACGGCCATGCTTCTCTGTGTCCCCTCTGGGCTGGAGCA  
 GGGGGCCTTGGGGGTCTGGGCAGACCTAGGGGTACTGCTGCCCCAAGACTGACTGT

2416 TCTGGGCTGGAGCAGGGGGCCTTGGGGGTCTGGGCAGACCTAGGGGTACTGCTGCCC  
 CCAAGACTGACTGTTAGCAAGTCCCAGACTGGATGCATCAGGTGAACCTCAGGCCAGCTTG  
 GGAATGAGTCCAGAGGGGCCCTGGGCCAESTGTGGCTCCTCTAGTTGTCTGTGCCACCT  
 CCTAGCAGCCCTTGGAGGAGCTGTCTGAAGCGCTCGCTGTGGGCTCCTCACCCGGGCTC  
 TGCAGGCAGCACTCACCTCTGGCAGTCACACTGTTTAGTACAAGCAAGTCCGAAGCTTC  
 [C, T]  
 GGCTCAGACAGGTTTGGTAAGGAGAGCAGAGCCACACACACTGGTCTTGGGTGGGCTGGG  
 GGAGTTCTGGGAGGGAGGTGGGTCCCAGTAGGGTATCCAACCTGCCTGCTTGGTCAGGG  
 CTGGCTCCGGTGACCGCACACTGGCAGTCCCTCTACTTGTGGGTTCGGGATGGGGACTT  
 GTTGCCTGACTGCCCTCTGCTGGTCTCTGAGCAGTTCTCCCCGAAGCCCCAGGACTGTT  
 GCCCTGTCTGAGCCTGTGAGGAAAAGAAGGGGTGTGAGGGAGCTGGACCCAGAGGAGC

4698 GCTAGGTGGCCCTGGGCTACACCAAGCCCTTCTGGTCTTGCCCCCGAGGTCTGGGGGT  
 CCGGAGACCCCATTAAGAATGGCCTGGGCCCCACAGGGAGCCACTGGGCCTGCTGCTGG  
 GGGGTCTGAATCCTGAAAGGAGAGCCTTGAGGAGCAGAGCCAGAGAGGCAGAGGCCCTTG  
 GGGCAGACACACCCCTGCCCTCTGGGGCCGCATGGAGACGGTGGTCTGTGCTGCTGAG  
 TCCTACACATGCATGTCTGCCCTGAGCATCCCCCAGGACAAGCCGCTCTGGAGTGGGTG  
 [A, G]  
 GGGTTTTATGCACCCTGAGGAGACTTTCAAGGCTTCTCTTGGGTGTTTCTGCAAAGTC  
 CTCCTCCCCTGCCCTCAAACCTGTGAGGGAAAAGGCCGCACTGGCCACCTGCTCCTCT  
 GGGCTGTGCGGGCCAGAGCCCAGAGGCCCAAGTTGGCTTCTGCCCACCTGCTGGCTTGT  
 GACCAT

5424 CCTCCTCATGACCCACAGGGTGAGCAGCCTGGCCTTCCCAGCCAGAGAACCCTCCTTCTG  
 GGGAGGCCCAGGGCGTCTCGGGGAGGGCAGTCTATTCTCCTCCCATGAGCCAGTGGAC  
 GTGTCTAGCAGGCAGCACCCCGGAGAGCCCTCCCACGTCTTCTCCATTGACAGGCCTT  
 TCCAGAGCGCAGGCGGGAGGGGCTGTGATTAGAAAAGAGTGAGGCTAGTGGCTTCTGGG  
 GAGGCACTGTGCCCAGGGGACAGTGCTGAGAGACAGCTGCCTCTACGCTGCCCTGTGCC  
 [C, T]  
 GGGGCTCCCGCTGCAATGCCCCCCTGTCTGCAAGTGAACGTGGGGCGACGGTGCAATGAGG  
 CCCTGCATGTGTGGCTCCACCTTGGGCGCCGAGAGCAGCTCTGTCTGGAGGGTGGTCAG  
 TGCATGTGGACAGAGCCCATGAGCTGTCTGGGTGACCAGCTAAGGGGACAAGGCAGA  
 GGCAGGGCTGAGAGGACACCCATCCTGCTAGGTGAGCCAGCTCAGCCATATCACACGG  
 CAGTGAGCATGGAGCTCAGTTCTTGCCAATGGCAGCTGAGTCTAGTACCATCCAGTCAG

8722 AAGGCCTGTGCTGGCCCCAGTCAGTGACAGAAGCGGCCCCAAGGCCAGGGCTGCTGGGC  
 AGCTCGGAATGAGGGCGAGCAGGGCTGCCCTTGGTGCCTGAGCCAAGGAGCCAATGGGAC  
 AGACCTCTGAGCCTGGGTGCCAAGTATGAGGTCTGAGACAGGGTGAGCGCCTGGGCTGGG  
 ACAAGGCCTCTGAGTGGGCGGCCAGCTGCAGCCACCCACCCCTACCCAGGAAGGCAG  
 GGCCCGGGAGGGCATGACCTCTGGGGTGTGGCTCAGCTGCCCCCACCACCTGACAC  
 [C, A]  
 GCTAGTCTGAGTTCCTCATCAGGGAGGAAGCAGCATCCTGCCTTCTCTAGGAAGAGCTT  
 GCATGTGGCCCAAGCCAAGGGGCTCCCCAGCACCCACGGGCATCTCTGGGTCTGGTC  
 AGAGGAGAAATCTGGATGCTTGCAGGAGCCCCAGGGTCATGGAGGAGGCTGGAGACAGGG  
 CTGTCTGGGGTGATGGGATGGCCCCCACCCTGCTCAGAGCCAGCCTGGGTGCTGGAAC  
 CACACTTGCCCTCAGGACCCTGGGCTTGTCTCTGGGAAAGAGTGGGGTCAGGCAAGGGG

9982 CCAGGAGTGTTTCAGGAAGTCAGTGAGGCAGAAGATACCCTCTCCCCACCAGGACCCACC  
 CTCAGCTCCTCCACCATCCTCAACAGGCCGACCCACAGACCACTCCGAAGGTCTGGCTTG  
 GTGGGGCTGGGCCAGGATCTGCAGGGGGAACAGCCCATAGTGGCACATTCCACGGCCCAT  
 GGGGAGACGGGGCCACGGTGGTGCAGTAGAGAGGTGTCTAAGCCAGTGGCAGCCAAAGGG  
 AGGGCTTGCCGTACCTCTGTGTTCCCTCAGTGCTGCTCTGTGGCTGCCTGAGAGGCAGG  
 [G, A]  
 CTTAGGGGCTCCCTGCCGGGAGGGGAGGGGTCCCCACCATGCTCCGCTCCAAGTGC

FIGURE 3, page 12 of 16

CCTCAGTGCCCCCTTGCCCTGGGGGCTCCTACAGGTGAACCTATAGCAGTACTCCCAAGG  
ATGTAAGAGTTGTGGCTGGTGGGTGCCGGCCTTCTGCTGGGGCGCTGTGCTGTGTCCCT  
CAGCTGTCTTAAGAGCTTGGGGCTTGTGGCCCGTAGGTCCCATATTTGCTGGAAGCA  
GGCTTGGTGTCCCCTGAGAACCCAGGCCAGGCTTCGGGAGCCAGCCCCAGACCGCCAC

10951 ACAGCAGCACCTCCGCCAGCCTGGACAGAGCTCCTGTCCATTCCATCCCTGCCGGCTGAC  
CCAGGCTCCTCCCCCAGCTGCTCCACGCCGCTCCATCCCTGTCCCCCACTCTGCTCTGC  
ACTTCTTTCTCGCAGGCTCTGGCCACCCACACCTCCTCTGTCTCCCTGTTCCTCTCTGG  
TGGTCTCCGCTTCTCTCTCTCACTTTCCTCTCTTCTCTCTCTGTGTCTTCTCTCTC  
TTCTGTAGGAGCCTCAAACACCGTCATCCATAACCCAGTGAGCGGATTAAGGTACTGC  
[C, T]  
CCACTTCTCTCTCCCGCTTCCCCAGGCAGGAGGCTCCAGGCCAGGAGAGAGGTCTGGG  
GCAGCATTTGTGCCAGAGTGAGAGGCAGATGTCCCATGGCCCTGGCCGCCCTCCCCGCA  
GTACGGTAGGGCCCCAGTCCGTCTTCGTGGGCAACAACAGGACAGACTGGCTCAGGCCCC  
AGGCGCGCCCTGGAGGTGCTTGGCACAGTTGCGCCCGGTCCCCATGTGGCCGACACTCT  
CAGACCAGGGCTCTGCGTGTCCACCTACGGCAGGCAGTAGGGCTTCTGAGGTCTGGAG

12603 AGTCTCTCTGCCAGGCTCATCTTGCTGGGAGAAAGTGGAGCCCTCATGTGTTGGGGATGCA  
GGGTGGCCACAGCACTAGGGTGGCAGGGCCGGCCTCGGACTCCGTGCCAGCCTGTGCTGG  
CTGCCGTGAGAATGCACCCCTGGTGAGGGGCGCCCTCCCAGGGACCAGCACAGAACTGGGT  
GTCTTCTCCGCTCACTGCCGCATGAGGTCCACAGAGCTGGGGCCCTGCAGCCGCCAGAGG  
GCATGTCCCTGAGCCCTGGCCTTTAAGCCCGTGGAAGCAGCCGAGGCAGAGATCAGC  
[T, C]  
TCAGAGCCTGGGCTGGTCTGACACAGGCCCAGCCCTGTCCACCTGCCCTCAGCCACGTC  
CCACCTATCCTTGGCCGCATCTGACCCGCTGCCTCCCGTGTTCCTCAGGAGTCTTCTG  
ACAGTGCCAAATACCACCATAGAGGATGAAGACGCTAAAGGTACCTGCACCTGAGTCTTG  
CCCCCCCAGCGCCTTGGCATTGCTGGGTGCTCTTTGAGGTGGGTGGGACTTGGGCAGG  
GTCAACTCTCTGCGACGCCTAGTTTATGCATGTGTTGAGGGGCTCAGGGACCCTGTAGC

14583 ACATCCTGAGCTCAGTGAGGAGGGGCTCGGGAGCCCCAGAAGCCGAGGGGCCCCCTGCCCT  
GCCCATCTCCGGCTCCCTTTAGCCCCCTGCCAGCCCCATGTAAGTAGCCTGGGTCTGCT  
GCTGTGGGGGTGATGTTGGAGGGCTGGCAACCCCTAGAGGGGCCACTCCAGAGCCGAGG  
GCAGGCTGAGCGTGGACCCCTGGCTCCAGCCTCATACCCCAATCCCTCACTGGGGCTT  
TCCAGGGTGGCCCCAGCCCATCGAGCCCCACCTCTTGTGAGGAGGGCCCTGGACCACTT  
[C, T]  
CCTGCTCAAGGCCACTGGGCAGGATGGGAGGCCCTGGAGGCTCGGGCCTCAATTCCAGTC  
TTCAGGGTCCGTGCAGGCCTCACTCCACCTCAGCTTGCAGGGCGGGGGGCTCCCTGCTAT  
TGAGGCAGGCTCTGATTGAGGGCTGATCCCAGGGCCCAAGGGGTCTAGAACACGGGACC  
CCTCCCACTGGCTCCTCCGCCTTGCCGCGCCTCGTGTGTCTGTCTGCCTCATGTTTAC  
GTCTCATCTGTTCACCCAGCCCCAGGGATCTCTGACATCCTGAACTCTGTGAGAAGG

17290 CTGTCCCTTGTGCCCCATCCCCACATCTGCCTCTGTGCCCCCTCAATCTCTGGCTTGGC  
TGCTGCCCCATGGTTTCTCTCCTGCGTGCCCCCGTGCCTGCCTTGTGTTACGTCTCGT  
CTGTTCCGCCCCAGCCCCAGGATCTCTGACATCTGAACTCTGTGAGGAGGGGCTCAGG  
GACCCCAAGAGCCGAGGGCCCCCTGCCAGTGGGGCCCCCGCCCTGCCCATCTCCGACTAT  
CCCTGGCCCCCTGCCACCCCATGTAAGTAGCACCTTGAGTGGCCGTGGCAGCGGCTGCC  
[T, C]  
GGAGGGGCTCGGGGCGTGCGAGCCTGGCAGTGGTCTCTGGGAAGGGCCATTCTTGCGGA  
GGAGGGCGGGGCACAGGATCCCTCTGCTGGGTCCCAGGGAATTGCTTTGAAGCACATGAA  
GGTGCCACTGGGTCTCAGAAAATGGAGGTTATGGTTATGAAGTGTGTATGACATATGTGT  
ATAGGAAGAGCGTCCGAAAGAGCAGGTTTGTGCGGACCCAGCATTCGCAACCCCTGAGG  
TCCACAGCTTTCTCTGATGGGAGGGGAATGGGTGGCAAAGGTCTGCGCGTGTGGCAAG

18188 ATCCCAGGGCTGCTGCCACCCACCTGTGGGGAGACACCAGACTGGGGGTGGTGTGGAG  
ATACTCTTAGAGAAGAGGCTGCTGGGCCACGGGCTCGGCATGGCAGGGCAGTGGCTAGGT  
AAGTACTTGAGGGACAGGTGGGGTCTGCTTGCCACCGTCCCCTCTGCAGGCTGGGCTGG  
GGGCTGCTGCAGGCGGCCAGGGCAGAAAGGTGTGGGGAGAGTGAACCCACAGGAGCAGCG  
GCTCGAGGAGGGGATGCAGGCTGCAGGCTCAAAGGGGCACTGGATCCACCTGGGTGCC  
[C, T]  
GAGAGAGCAGGGGGCAGCCCCCTGGAGGGGTACTACCCCCAGAGCTTCTGTGGTCCGGCTG  
AGGACCCCCAGCAGGGGTTGACTGAGGGGATCAGAGGCAAGCAGCTGAGGGGAGAGGCCA  
GGTTCTTGATGCTGATAGGGTCCGGGTGCTGGGCGACCAGAACTCAAGGAGGGAGGCAT  
GGGGAGGGGCCCGCTGCAGCTGGGGTGGGTGCACCGCAGAGCCTCTGGGAGTGGTCAGA

FIGURE 3, page 13 of 16

ACCCCCGACACCTGCCACTTCTACAGCAGCTCATCTGATTTTAAGGGGCTTGCTGCCCTT

19911 AGCACGGTTACCACTCTTGATTGGAACCTGACCATGCATCTCCTCTTCTGTTTACTTCA  
CGCTTTCTCTTCCCATCAACTCCCATTTTAATTACAATTTGTTTAAAAGCACTGCATATT  
ACTTCATTAAACAGAAGATTAGTTTCACTTACCATTAGTGTAAGGTGACTATAGAACCAA  
AGCAGACTGGAAACCAAATGACATAATGTCATTCTCTCTCCATTCCAGCTGCCTGCTGC  
TGTGCGCCTGAGAACCCTGTGGAGTGGGAGGGGAGCTGTCTCTGTACATTAGAAAGGG  
[A, G]  
GGTTAACTAAGTGACAGGAGGTGTTTGGGACATGTGGACACCAGACTTCTCTCTTGATGC  
AAGGAGGGCAGAGCCAGGCAGCCTAGTGGGGGCTGGCTTGGGGGCTGCTGGAAGGACTGG  
CTACAGGTGGAAGAGAGGTGACACCTGAAGCTTGGGGCCACCTCCAGGAAAGGACAGGTG  
AAAGTGGAGGCATGAGGCAGGGGACAGGCAGGTGCCAGGCAGAGGTGGAGAGGAGGCAG  
GAACATAGCAGCTGGGGCGGGGCGGGCCCTCAAGTGTATATGCTACTTCTCTGGGGCC

21328 GCTGGGCACAGTGGCTCATACCTGTAATCCCAGCACTTGGGAGGCCGAGGTGGGCAGAT  
CACTTGAGGTTAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAACCTCATCTCCACTA  
AAAATATACACACAAAAATTAGCTGGGTGTGGTGGTGTGCACCTGTAGTTCCAGCTAC  
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCTGGGAGTCAGAGACTGCAGTGAGCCGA  
GATCATGTCACTGCACTCCAGCCCCGGTGACAGAGTGAGACTCCATCTAAAAA  
[C, A, G]  
AATTCCTCTCTGGGAATTTAGACCACAGACAGGTGTCATGTATGTGGCCGTGGAGGC  
AGCACTCACAGCAAAGAGTGGAAACGTCAACACAGGGCCTGCCTTCTGGTGAAAATGGTG  
TCCTGCAGGGCGGGCAGCTGTTTGGGGCAGGTGTCCAGGTGCGGCCCTGCAGCAGCCTG  
AGGGTCACAGAGCGCAGTGTGGGAGTGCAGAGACTTCCCCACAGGGAGAGTTCCAGG  
AACCTGCTTCCGGTGCACTTCTGGGGGTTGAGTTTTTCCACGACGAATTACTTTGAG

21391 TTGAGGTTAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAACCTCATCTCCACTAAAA  
ATATACACACAAAAATTAGCTGGGTGTGGTGGTGTGCACCTGTAGTTCCAGCTACTCG  
GGAGGCTGAGGCAGGAGAATCGCTTGAACCTGGGAGTCAGAGACTGCAGTGAGCCGAGAT  
CATGTCACTGCACTCCAGCCCCGGTGACAGAGTGAGACTCCATCTAAAAA  
TTCCCTCTCTGGGAATTTAGACCACAGACAGGTGTCATGTATGTGGCCGTGGAGGCAG  
[T, C]  
ACTCACAGCAAAGAGTGGAAACGTCAACACAGGGCCTGCCTTCTGGTGAAAATGGTGTCC  
TGCAGGGCGGGCAGCTGTTTGGAGGCAGGTGTCCAGGTGCGGCCCTGCAGCAGCCTGAGG  
GTCACAGAGCGCAGTGTGGGAGTGACAGAGACTTCCCCACAGGGAGAGTTCCAGGAAC  
CTGCTTCCGGTGCACTTCTGGGGGTTGAGTTTTTCCACGACGAATTACTTTGAGAAA  
CCACTGTTACTCGTGTGTATAGGTGAGCGTGGCTGTGCATGTGTCTGTGTGTGAGTG

22588 GCTGCTTCCTCTCCCCGGCCTCCGGGTGGCCTTGCTGACGGCTCCTTCTCTGAGGCAGG  
TCTCTGCCTTCTCGCCTGGTGCCTGCACTCAGTAGCCCCCTACCAGAGCTGCTGGGTGA  
AGGAAGCACTAAGAACCCAGGCTCGGGAGGAGAGTGGGGCCGGAAGCTGCAGGGAAGC  
GCAGGGCCAGGCCTGGTGGGCCAGGGGCTGGCTACGGGAGGGCAGGAGGGAGACTGTG  
GCGGACAGCACGTGGGGCCAGGAGGTGACCTCCAAGTGGATTGTGGGTGGGTTTTTGTCT  
[C, T]  
TCTTCTGCATTTTCCAGGCATTTTGTAAATGTGGATAGAATATTTCTGTTCTTCAAAAAT  
ACTTTAGTTAAGAAAAATAAGATGGAAGCTGTTGCACTTGAAAATGAGGAAGCCACTGGT  
GATGCAGGGGGGGCGGCGGAGAGGACCTTCTTGCAAATAGCGGCAGGAACACGGCATGG  
ATGCAGCTCGCGTCCCCCAGGCCCTCCCCCTGGGCTGTGTGGAGGGGTCCGGGGGAATG  
GGCCAGCGCCAGTGGTCACTGGCCATGTCTCCCCACAGCCCGGAAGCAGGAGATCATT

22965 ATAAGATGGAAGCTGTTGCACTTGAAAATGAGGAAGCCACTGGTGATGCAGGGGGGGCGG  
CGGAGAGGACCTCTTCTGCAAATAGCGGCAGGAACACGGCATGGATGCAGCTCGCGCTCC  
CCCAGGCCCTCCCCCTGGGCTGTGTGGAGGGGTCCGGGGGAATGGGCCAGCGCCAGTGG  
TCACCTGGCCATGTCTCCCCACAGCCCGGAAGCAGGAGATCATTAAGACCACGGAGCAGC  
TCATCGAGGCCGTCAACAACGGTGACTTTGAGGCCTACGCGTGAGTCCCTGGGGCTGGGG  
[-, G]  
GGGGCTGTGCAGGACAAGGATGTGGGACCCCTTGGGGGGGCTGCTCAGAGTCAGGGGTCC  
ACGGGGCCCCCTCCTCACTTGGATTTGGCCCCCAGGAAAATCTGTGACCCAGGGGTGACCT  
CGTTTGAGCCTGAAGCACTGGGCAACCTGGTTGAAGGGATGGACTTCCACAGATTCTACT  
TCGAGAACCGTGAGTGAGGAAGCCCGGTGGGCATGAGGGGGCGGTGCCCCCAGGAGAGC  
CTCTCGGCCCTCCAGGGACAGCATGGTGGCTGCCTATGGAAGCCCTGTCCCCTCTGTG

23498 CCCGCCAGAGGCCATACCCAGCCCCCAGAATCCCCTCTTGAGGGGCCCATGCTGCTCC

FIGURE 3, page 14 of 16

CAGGAGAGCCGAGCCTCCCCAATAAGGGGAGTTGAGAGAGGGAAAGGATTAGGCTGGTGG  
GGTGGAAAGACGGGCACCAGGGCAGTCATGGTAACCCGAGACCCCCGCCCCGCTGCTGTC  
CACAGTGCTGGCCAAGAACAGCAAGCC  
[G, A]  
ATCCACACGACCATCCTGAACCCACACGTGCACGTCAATTGGAGAGGATGCCGCTGCATC  
GCTTACATCCGGCTCACGCAGTACATTGACGGGCAGGGCCGGCCCCGACCAGCCAGTCT  
GAGGAGACCCGCGTGTGGCACCCGCGACGGCAAGTGGCAGAACGTGCACTTCCAATG  
TCGGGCGCGCCTGTGGCCCCGCTGCAG

23663 GCCTCCCCAATAAGGGGAGTTGAGAGAGGGAAAGGATTAGGCTGGTGGGTGGAAGACGG  
GCACCAGGGCAGTCATGGTAACCCGAGACCCCCGCCCCGCTGCTGTCCACAGTGTGGC  
CAAGAACAGCAAGCCGATCCACACGACCATCCTGAACCCACACGTGCACGTCAATTGGAGA  
GGATGCCGCTGCATCGCTTACATCCGGCTCACGCAGTACATTGACGGGCAGGGCCGGCC  
CCGCACAGCCAGTCTGAGGAGACCCGCGTGTGGCACCCGCGACGGCAAGTGGCAGAA  
[T, C]  
GTGCACTTCCACTGCTCGGGCGCGCCTGTGGCCCCGCTGCAGTGAAGGTGAGTGTCTGT  
GCTAAGTGACAGCTGGGGCAGAGGGGTGGCGGTGGTGTGAGTGGCTGCAGCCTGGGGAGG  
CGATGGGGAGCGGTGGGGCCTGTGGCAGAGCCCATGCCTGGGAAGTCCCTGAGCTTTCCT  
GGTGAGGCCACAGGAATGATGTCAAATTAGGGACCACGGCAGGCTGGGTGTGGCAGGCC  
CCCCAGAGGACTGGGGAGCTGGTGAGGGCCTGAGCAGTCCACACTGGCCAGAGCTGGGTG

25427 TGTGGCAAGAGGACTCTGCCTGGGCTGGCCCCCTCCTGTGTGAGGTGTCTGTCCCTTCT  
CTGCTGGCCAGCAGCAGATGCACTGGCAGCTCCCAACCCTGTTTCCGCCCCCTCGGCCCTC  
CCCCAGCCTGTTTCGGCTTCTCTGCAGCCCCGCAAGGGGGAGCAGACTTTTGACAAAGGACT  
GCGGGCCTCGCTCAAGTCCCTGAGCCCCAGCTGAAGCTGGGAGGGGAGGCCAGGCTTTG  
TGTCTGGGCATATTCTGCTGCTGATGGGGTTTGGGAAGCCTGGGGCTTGGGGTTTGGTC  
[A, G]  
GGTGGTGACAGCTAGTGGCAGAGCGGGATCAGAGGTGGTGGCTGCCAGCTTCTGGGCTGA  
GACAAGGGTCTGTGCAGGGGTACTGAAGTGGGAGTGCCTTTGGAATCTGGGCCGGGAG  
CAGAAGGGAGCAAAGCTACAGTGGGAGCCAGCCTAGGGCACATGGGAGGCGTGAGGGCA  
GTGCTGCCCGTGCACTGTGAGGTGTGCCAGTGCCTTGGCGGGCTGCAGTGCCTGTGAGGG  
CACCTTCTAGGTGGGCCAGGGATGCAGCTATGGAGATAAGGCGGGCTGGGGACAGAAACA

27727 GCAAACCTCTTAGGTTGGAGTAAGGAGTAACCCCTGCCAAGTTTCTCCTGTCTCAGGCT  
CCACCCACCACCTATGCTGCCTGGCCCCATGGGGCACACGCTCAGGCCCAGCCTGGGAAA  
GCAACTGCACCTGCCTGTGCTATGCTGGCCCTTCTCAGCCTCAATGCCCTCCTCCCTCCC  
CGACGCACCCCTCGTGGCCCCGCTGGGCCCCCTGATGCACCCCTCATGTCTCCATGGCAAC  
CTGCTCAGAGTGTGGCCCTGCCCTTGGCTCCCCTCCACACCTGTGTCCAGGCAGTGCCA  
[C, T]  
GGCACTTTCCTAAACAGAAGGATGGGCTTCAAACAGTCCCAGACACTAAACACACCTGC  
ATTTTGGGTCCAAGTAACCTCTGACAAGACGAGTGCCCTACACACCCTCAGTCTCTATCC  
ACTATGGGCAAGGAGCCTGAAGGATCCCCCAGAACTGGCTAAAGCCCTCAGTCTCCTCCT  
CCACCCTGAGCACCTTCACGCGCAGAGTGGCCCTGGATGTGAGCTTCTTGCTCCCCATG  
GTCTGCACCTGGACAGGTGCTCTCAGGTGTGTGGGTGGGCAGGTGGCAGGTCCCAAGAGC

27834 CCAGCCTGGGAAAGCAACTGCACCTGCCTGTGCTATGCTGGCCCTTCTCAGCCTCAATGC  
CCTCCTCCCTCCCCGACGCACCCTCGTGGCCCCGCTGGGCCCCCTGATGCACCCCTCATG  
TCTCCATGGCAACCTGCTCAGAGTGTGGCCCTGCCCTTGGCTCCCCTCCACACCTGTGTC  
CCAGGCAGTGCCACGGCACTTTCTAAACAGAAGGATGGGCTTCAAAAAGTCCCAGACA  
CTAAACACACCTGCATTTTGGGTCCAAGTAACCTCTGACAAGACGAGTGCCCCCTACACAC  
[T, C]  
CTCAGTCTCTATCCACTATGGGCAAGGAGCCTGAAGGATCCCCCAGAACTGGCTAAAGCCC  
TCAGTCTCCTCCTCCACCCTGAGCACCTTCACGCGGCAGAGTGGCCCTGGATGTGAGCTT  
CTTGCTCCCCATGGTCTGCACCTGGACAGGTGCTCTCAGGTGTGTGGGTGGGCAGGTGGC  
AGGTCCCAAGAGCCAGGTGCAAAGAATCTAGGCCAGTGGCCACGAGTGTGAGTGTGCTG  
TCCCCAGCATGGTATCTAGGGCTCCACTTGCCTATCAGCTGTAATCGGAGGAGGCTTTC

AGAATCTAGGCCAGTGCCACGAGTGCTGCAGTGTCTGTCCCCAGCATGGTATCTAGGG  
CTCCACTTGCCATATCAGCTGTAATCGGAGGAGGCTTCCAGGCCAGGCCCTCCCCAGGAA  
GGCTGCAGGCACTGCGGATCGTGCGCCCTCACATGATTATCCTGAGGCCCTTCTGCAG  
ATGCCATCAGGGCAGCACTCTGATGAGGTATTAGGGCACAGCACAGGGCTAAGCCAC  
CCTGTACTGGGCCAAGCGCTACAGGCAAAAAGGACACCACCGACGGGCATTTCAATCATC  
[G,A]  
CTTTTATTTTTATATATTTTTGAGAGGGAGCCTCACTCTGTGCCCCAGGCTGGAGTGCAG  
TGGCGCGATCTTGGCTCACTGCAACTTCTCCCTCCTGGGTTT

[illegible]